VOLUME XXVIII JANUARY, 1938 PART 3

THE PROBLEM OF RETROACTIVE INHIBITION IN RELATION TO RECOGNITION

BY O. L. ZANGWILL

Cambridge Psychological Laboratory

I. *Introduction* (pp. 229–233).
II. *Experimental method* (p. 233).
III. *Experimental procedure* (pp. 233–236):
(1) *The exposure series* (pp. 233–234):
(*a*) *Material* (p. 233).
(*b*) *Time of exposure* (p. 233).
(*c*) *Instructions* (pp. 233–234).
(2) *The interpolated activity* (p. 234):
(*a*) *Experimental group* (p. 234).
(*b*) *Control group* (p. 234).
(*c*) *General note on technique* (p. 234).
(3) *The identification series* (pp. 234–236):
(*a*) *Material* (p. 234).
(*b*) *Time of exposure* (p. 234).
(*c*) *Text of instructions* (pp. 234–236).
(*d*) *Mode of recording responses* (p. 236).
IV. *Subjects* (p. 236).
V. *Mode of analysis of results* (pp. 237–238):
(1) *Accuracy of identification* (pp. 237–238).
(2) *Certainty in correct identification* (p. 238).
(3) *Immediacy of correct identification* (p. 238).
VI. *Results* (pp. 239–243):
(1) *Accuracy of identification* (pp. 239–241).
(2) *Certainty in correct identification* (p. 241).
(3) *Immediacy of correct identification* (pp. 241–242).
(4) *Deviations from accuracy in recognition and recall* (pp. 242–243).
VII. *Discussion* (pp. 243–245).
VIII. *Conclusions* (pp. 245–246).
References (pp. 246–247).

I. INTRODUCTION

ONE fact has long been known which appears to conflict with the notion of an underlying similarity between the processes of recall and recognition. G. E. Müller[1] is stated to have been the first to report that retro-

[1] *Vide* Hollingworth (6).

active inhibition fails to appear when partially learnt material, the learning of which has been followed by some different task, is merely recognized. If this be indeed the case, it constitutes an important point of difference between the general characteristics of recall and those of recognition.

Müller's conclusion was confirmed as a result of systematic investigation by Rose Heine (5). Heine employed nonsense-syllables as experimental material, and the conditions were so arranged that had 'retention' been tested by recall, very considerable retroactive inhibition would most certainly have been manifested. Testing by recognition, she was unable to discover any indication of such an effect, and was obliged to conclude that interpolation of a new task between presentation and re-presentation reduces neither the possibility nor the accuracy of subsequent identification of the original material.

This apparent failure of an interpolated task to produce any deterioration in identification assumes major theoretical importance in connexion with the comparatively recent work of von Restorff (12). This investigator, in the course of her study of the influence of 'context' upon recall, observed that an element originally presented in a context of generically *dissimilar* elements is more frequently and more accurately recalled than is a corresponding element originally presented as a unit in a group of generically *similar* elements. She was likewise able to demonstrate an analogous, though less marked, positive effect of 'isolation' upon recognition. A further series of experiments, bearing specifically upon the question of retroactive inhibition, established that the decrement in recall performance attributable to this factor is more marked when the interpolated task consists in learning items similar to those originally presented than when it consists in learning items dissimilar to those involved in the original activity.

In discussing these results, Koffka (8) concludes that "the nature of retroactive inhibition has been clarified" and maintains that "it results from the aggregation of similar traces". It must, however, be pointed out that von Restorff repeated this last experiment with the sole modification that recognition was tested in place of recall. Under these conditions, however, no difference could be detected between the results from the experimental and from the control group. She was forced to conclude, as was Heine before her, that an interpolated task, however similar or dissimilar to the original activity, has no significant detrimental effect upon the ability to recognize items involved in the original activity.

Any satisfactory interpretation of von Restorff's principal result

should obviously be capable of explaining also the negative outcome of these experiments upon retroactive inhibition in relation to recognition. Von Restorff herself attempts an interpretation along the following lines: In the first place, she argues, the activity involved in recognizing is less "high-grade" (*spezifisch*) than is the case in literal recall. Accurate reproduction necessitates the operation of past experience in a form more highly specific and individualized than is required for the mediation of recognition. Interpolated presentation of similar material, which reduces the 'individuation' of past experiences, must consequently interfere with reproduction very much more seriously than with recognition. Accurate identification of previously presented material may therefore occur even after introduction of conditions known to reduce the possibility of accurate reproduction. Von Restorff further suggests that the occurrence of 'generic' recognition may complicate the issue. She points out that if material similar to that originally presented is utilized in the interpolated task, it may actually result in an *increase* in facility in subsequent identification. This she believes to be due to the fact that 'trace-aggregation' produces conditions favourable to the occurrence of a 'generic' type of recognition. In view of our ignorance concerning the nature of the relationship between 'generic' and 'specific' types of recognition, suggestions of this type she believes to be permissible.

Von Restorff's suggestions, though extremely cautious, can scarcely be regarded as furnishing an adequate explanatory hypothesis. Koffka, in any case, makes no mention of such considerations, but simply states that "it seems plausible to assume that repetition after an interval (the conditions of retroactive inhibition) will produce an effect on recognition too small to be detected by our present methods".

If Koffka's assumption be correct, it would seem obvious that the next step should consist in an attempt to devise an adequate method for the resolution of this problem. Such an attempt constitutes the essential purpose of the present investigation.

There is a further experiment which should be discussed, since Koffka believes it to furnish support for his theory that retroactive inhibition does in truth have an effect upon recognition. This is a comparative study by Dahl(2) of the influence of identical periods of sleep and waking life respectively upon recognition. Dahl's general conclusion is that accuracy in recognition is favoured by sleep. He stipulates, however, that owing to inevitable difficulties in the experimental technique, this conclusion must be accepted with caution.

Experiments essentially similar to that of Dahl, but in which recall

was tested in place of recognition, had previously been carried out by Heine and by Jenkins & Dallenbach (7). These workers concluded from their results that sleep is favourable to 'retention', and suggested that retroactive inhibition is the agency responsible for the difference in performance in recall as between intervals filled by sleep and by ordinary waking activity. In discussing these results, Dahl however argues that if recognition be not affected by retroactive inhibition, an explanation of this character would not account for his own results.[1] Koffka nevertheless maintains that "after Restorff's results, it seems plausible to interpret Dahl's results as effects of retroactive inhibition".

Koffka's interpretation is strengthened by the results of some very recent experiments. Lund (9), in his work upon criteria of confidence in recognition, had suggested that certain of his findings could be satisfactorily interpreted on the assumption that recognition is susceptible in some degree to the influence of retroactive inhibition. Gibson (3), employing Lund's original exposure material, carried out a more definite experiment in which nonsense words originally presented under specified conditions were subsequently re-presented together with a number of new items. These differed from the originals only in that one or two letters were changed as compared with the corresponding words in the exposure series. According to Gibson, this technique renders the original items harder to discriminate from their context, and constitutes in consequence a more exacting test of recognition. In the case of the experimental group, the task interpolated between presentation and re-presentation consisted in learning a further series of nonsense words, whilst the members of the control group were instructed to relax. Under these conditions, Gibson reports a significantly greater number of errors in recognition from the experimental group as compared with the control group. Though believing his results to indicate that "a significant amount of retroactive inhibition may appear when retention is tested by the method of recognition", Gibson expressly stipulates that further experimental work is required before the issue can be finally decided.

An attempt to repeat the original experiments of Heine has recently been undertaken by McKinney (10). The outcome of this study is favourable to the hypothesis that retroactive inhibition is not without influence upon recognition. McKinney indeed claims that his results show "small but consistent amounts of retroactive inhibition with all material used". It is incidentally necessary to point out that not all these results are

[1] For a more comprehensive discussion of these questions, see van Ormer (11).

statistically significant, a point which is however explicitly admitted by McKinney himself.[1]

The outcome of the present investigation may enable us, it is hoped, to interpret more adequately the various inconsistencies in the earlier experimental work, and thereby to contribute to a fuller understanding of the nature of the relationship between recognition and recall.

II. EXPERIMENTAL METHOD

The experimental procedure was briefly as follows: A number of simple figures were presented to the subjects under controlled conditions. The subjects were then divided into two groups. The members comprising one of these groups then undertook an activity involving a minimum of concentration, whilst the members of the second group were instructed to learn and reproduce a further series of rather more complicated figures. Both groups were then shown a final series of figures, in each of which was incorporated *one* member of the original series. As a result of incorporation within more comprehensive drawings, the original figures were obviously rendered less readily identifiable. Under such conditions, the 'task' still remains one of identification, but is evidently more difficult and more exacting than is normally the case in everyday recognition.

If we should find a statistically significant difference in accuracy of identification as between the two groups, we may justifiably conclude that recognition is under these conditions subject to the detrimental influence of retroactive inhibition.

III. EXPERIMENTAL PROCEDURE

(1) *The exposure series*

(*a*) *Material.* Eight simple figures (E-figures) were drawn on cards suitable for exposure in the epidiascope. The selection of suitable drawings involved a considerable amount of preliminary investigation, the results of which will be here omitted. The first eight figures shown in Fig. 1 (col. 1) were finally selected. The ninth figure had been employed in place of No. 8 in a small number of cases, but was discarded on account of the possibility of confusion with No. 1 in the recognition series. This substitution has, of course, been taken into account in computing the final results. The order in which the cards were presented was the same for all subject-groups.

(*b*) *Time of exposure.* Each figure was projected for 3–5 sec. upon a screen by means of an epidiascope. The interval between exposure of successive cards in the series was fixed at 15–20 sec.

(*c*) *Instructions.* The subjects were given the following oral instructions: "A series

[1] The experiment reported in this paper had been planned and performed before the attention of the writer was called to McKinney's findings.

of simple drawings will now be shown, each for 3–5 sec. Please look at each successive drawing carefully, with a view to recognizing it later."

(2) *The interpolated activity*

(*a*) *Experimental group.* A series of eight more complicated figures were exposed in the same manner as the preceding series. Each figure was projected for 3–5 sec., and a second presentation of the entire series, under the same conditions and in identical order, followed directly upon exposure of the last member. After termination of this second exposure the subjects were instructed to attempt to reproduce each member of the series, if possible in the original order. The time allowed for the whole procedure—presentation and reproduction—amounted to 10–12 min.

(*b*) *Control group.* The members of this group were instructed to relax generally and to listen to the experimenter, who gave a short and informal talk on the significance of control groups in experimental psychology. This talk occupied a length of time equivalent to that spent by the experimental group in carrying out the total interpolated task.

(*c*) *General note on technique.* Owing to the practical impossibility of obtaining the requisite number of subjects in two large groups, the experiment was perforce carried out by stages, employing a number of smaller subject-groups. Great care was, however, exercised throughout to obtain identity of time-relations between the various stages of the experiment as between all experimental and their corresponding control groups.

(3) *The identification series*

(*a*) *Material.* The subjects were presented with a series of eight relatively complicated figures (I-figures). In each one of the latter was incorporated *one* of the figures originally shown in the exposure series. These original figures were so incorporated as to preserve their size and relative orientation in space unaltered. The order in which the figures were presented differed from that adopted in the E-series, but for purposes of comparison they are shown in Fig. 1 (col. 2) in the order corresponding to the arrangement of E-figures in col. 1.

The object of such incorporation was, of course, to render the original figures less readily identifiable. A certain amount of preliminary work was necessary in order to obtain figures of moderate, and approximately comparable, difficulty. In the selection of such material, much use was made of the Gestalt principles of 'organization', and in particular, of the findings of Gottschaldt(4). Certain of the figures used were, in fact, adapted from the published work of this investigator.

(*b*) *Time of exposure.* Each member of the identification series was projected for 15 sec. After the termination of exposure of each figure, the subjects were instructed to respond in a prescribed manner, which is described in the following section.

(*c*) *Text of instructions.* The instructions were given in a somewhat informal manner, with a fictional illustration by the experimenter of the type of response requested from the subjects. Great care was nevertheless exercised in preserving the essential text of the instructions standardized in the case of all subject-groups. The instructions were the following: "You will now be shown a further series of figures. Certainly in some of these, perhaps in all, figures which occurred as members of the *original* series are incorporated. These original figures are so incorporated that their

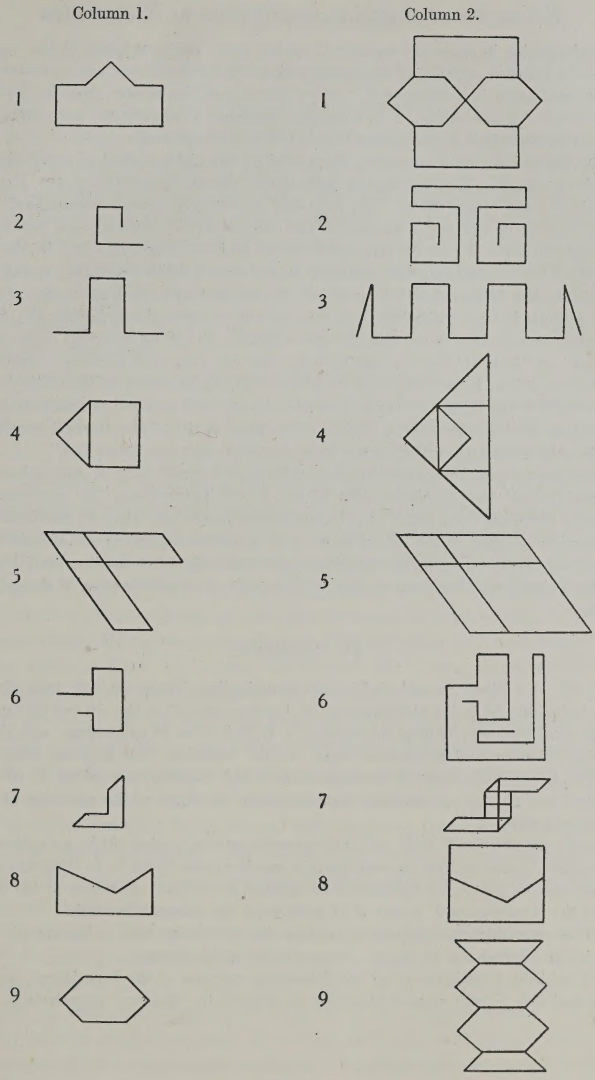


Fig. 1.

size and position in space are unaltered. At the most, only *one* figure of the original series is included in each one of the figures which will now be shown, and no member of the original series is incorporated in *more than one* of the present figures. You are requested to try to *identify* (i.e. to recognize) members of the original series that may re-occur incorporated in the figures which will now be shown."

(*d*) *Mode of recording responses.* Each subject was given a sheet of paper divided into three columns. The headings of these three columns were respectively 'Figure', 'Certainty', and 'Immediacy'. The following instructions were communicated: "If you succeed in identifying a member of the original series occurring in a member of the series which will now be presented, please make a rough drawing, in the first column, of the *original* figure in question. In the second column indicate, in any way you please, the degree of certainty which you experience in each such case. In the third column indicate whether such recognition occurred immediately the figure appeared, or only after some thought and analysis. It will be sufficient if you write 'at once' or 'delayed', as the case may be, but use your own modes of expression wherever possible. If you believe *no* member of the original series to be present in any one or more of the figures which will now be shown, or if you find yourself unable to identify any such member of the original series, please write '*None*' in the first column. Number the figures according to the order in which they are presented."

Answers to questions posed by the subjects were given only in cases where the former involved further clarification of the above instructions. The experimenter expressly refrained from amplifying these instructions in any way. In particular, no information relating to the actual number of figures constituting the identification series was imparted, and the experimenter invariably refused to commit himself further if asked whether *every* member of this series incorporated one of the original figures.

IV. Subjects

In all, 71 subjects co-operated in the investigation. These subjects were divided so as to include 36 in the experimental (E-) group, and 35 in the control (C-) group. Of the total subjects, 32 were members of a W.E.A. class in psychology, and the remaining 39 were undergraduates from various faculties. The E-group comprised 17 W.E.A. students and 19 undergraduates. The C-group comprised 15 W.E.A. students and 20 undergraduates. Approximately one-third of the members of each group were women.

In the case of the W.E.A. unit, the experiment was carried out in a single stage; the 'control' group of 15 being addressed in another room by Mr R. C. Oldfield (whom I have also to thank for permission to employ as subjects members of his class), whilst the 'experimental' group of 17 performed the interpolated task.

In the case of the undergraduate groups, the experiment was, as has already been mentioned, carried out by stages, using smaller subject-groups.

All subjects were ignorant of the theoretical purpose of the experiment, and the great majority totally unused to acting as subjects in 'Memory' experiments.

V. MODE OF ANALYSIS OF RESULTS

(1) *Accuracy of identification*

Four categories were arbitrarily defined by the experimenter, into which the various types of response could be grouped. The differential criteria determining such grouping uniformly depended upon the degree of accuracy in identification as expressed by the degree of accuracy manifested in reproduction of the original figure. Accuracy, for present purposes, referred only to general structure and arrangement of detail in the figure. Proportional alterations in shape and size of the reproduction as compared with the original were ignored, as were also inadequacies in technical execution.

These four categories may be defined as follows:

Category 1. This category included all cases in which a member of the E-series, now occurring incorporated in a member of the I-series, was correctly drawn immediately after termination of exposure of the latter. By 'correctly', it is implied that the general structure and detail of the original figure were clearly reproduced. In the case of some of the figures, however, certain quite specific and constantly recurring types of deviation from accuracy were nevertheless permitted. The theoretical significance of such types of deviation is discussed below; for the present it is sufficient to point out that great care was exercised to secure, in the case of each figure, definite criteria determining inclusion within this category. Such criteria were, of course, employed with scrupulous equality in classifying the reproductions of the two subject-groups.

Category 2. In this category were placed all cases in which, under the conditions stated in the preceding definition, a reproduction of a member of the E-series manifested significant deviation from the original as regards general structure and/or important detail.

It is obvious that the division between categories 1 and 2 must necessarily be rather arbitrary. Enormous care was nevertheless expended in classifying reproductions from the E-group and from the C-group with absolute impartiality.

Category 3. This category included all cases in which a reproduction did not correspond with that member of the E-series actually incorporated in the member of the I-series which had just been exposed, but was intended to represent some other member of the original series. The possibility of this occurrence was reduced as far as possible by careful initial selection of the experimental material. It nevertheless occurred in a small number of cases.

Category 4. This category included all cases in which (*a*) the subject wrote "None", (*b*) the subject produced no drawing and left the corresponding space blank, and (*c*) a drawing bore absolutely no relation to any figure which had appeared in the original E-series. The (*c*) type of total failure was rare, and where it occurred, appeared due either to an importation from the interpolated series or else to mere constructive guesswork.

(2) *Certainty in correct identification*

Certainty was taken into account only in the case of completely correct identifications, viz. only in the case of identifications which could be placed in category 1 as regards accuracy.

Four grades of certainty were defined by the experimenter:

Grade A. The following verbal reports were taken as indicative of complete certainty: 'certain', 'definite', 'positive', 'sure', 'quite (certain)', 'yes'.

Grade B. The following reports were taken as indicative of less complete certainty: 'fairly certain', 'fairly sure', 'fair (certainty)', 'not quite sure', 'not quite certain', 'probably', 'very probably'.

Grade C. The following reports were taken as denoting uncertainty: 'uncertain', 'rather uncertain', 'not certain', 'not at all certain', 'doubtful', 'very uncertain', 'not very sure', 'no (certainty)', '(certainty) none'.

Grade D. Cases of correct identification in which the degree of certainty experienced was not recorded by the subject.

(3) *Immediacy of correct identification*

As in the case of certainty, the question of the 'immediacy' of identification was taken into account only in the case of identifications classed in category 1.

Two categories of immediacy were employed:

Category 1, '*immediate*'. The following reports were taken as indicative of immediate, or relatively immediate, identification: 'immediate(ly)', 'almost immediate(ly)', 'more or less immediately', 'at once', 'nearly at once', 'yes (i.e. immediate)', 'soon', 'fairly quickly', 'no delay', 'instantaneously', 'instantly', 'rapidly', 'after very little thought'. In cases in which approximate times of identification were stated, times of 5 sec. and under were taken as indicating 'immediate' identification.

Category 2, '*delayed*'. The following reports were taken as indicating relatively delayed, and probably indirect, identification: 'slow(ly)', 'not immediate(ly)', 'very slow', 'no (immediacy)', 'delay(ed)', 'after (some) thought', 'long time', 'after some time', 'with difficulty'. In cases in which approximate times of identification were stated, times of over 5 sec. were taken as indicating 'delayed' identification.

VI. RESULTS

(1) *Accuracy of identification*

In Table I is set out: (*a*) the percentage of identifications falling within each of the four prescribed categories respectively, in the case of each subject-group, and (*b*) the mean value for the identifications falling within each given category per subject-group.

It will be seen that the percentage of identifications falling within category 1 is 14·78 % less in the case of the E-group than in the case of the C-group, whilst the difference between the means of the C- and the E-group amounts to 1·18. Further, the percentage value of identifications falling within category 4 is 12·65 % greater in the case of the E-group than in the case of the C-group, the difference between the corresponding means being 1·02.

Table I

	Category							
	1		2		3		4	
Subject group	%	Mean	%	Mean	%	Mean	%	Mean
C	57·14	4·57	7·50	0·60	3·57	0·29	31·79	2·54
E	42·36	3·39	7·99	0·64	5·21	0·42	44·44	3·56

The significance of these differences between the means for the two groups has been worked out by means of Fisher's t method. For category 1, a value of $t=2{\cdot}7891$ was obtained, whilst for category 4, the value is $t=2{\cdot}4836$. For $n=69$, these two values of t may be regarded as significant. In the case of the first category, at least, the significance may be regarded as unequivocal. We may therefore conclude that the introduction of a learning activity immediately after presentation of simple material has a detrimental influence upon the ability to recognize such material when subsequently re-presented under the conditions of the experiment. Such detrimental influence is demonstrated (*a*) by a smaller percentage of correct identifications as compared with the corresponding value obtained for the control group, and (*b*) by a higher percentage of total failures in identification as compared with the corresponding value obtained for the control group.

The values for categories 2 and 3 were not subjected to a statistical analysis. As will be seen from the table, the figures in category 2 for the two groups are very similar, and the difference in percentage value between the two groups in category 3 amounts only to 1·64.

The next relevant question is the following: Are the aforementioned

differences between the results for the two groups due to a fairly uniform detrimental effect of the interpolated task upon the identification of each individual figure, or do they depend upon a very marked effect on the identification of one or two members of the original series only? This question is answered by reference to Tables II and III. Table II shows the actual distribution of identification responses in the case of every figure in the exposure series for the two subject-groups. The means recorded at the foot of each column represent percentage means, i.e. values obtained after correction for the different 'weighting' of the percentage values for Fig. 1, nos. 8 and 9 as compared with the remainder. In Table III are

Table II

Fig. 1, no.	C-group Category				E-group Category			
	1	2	3	4	1	2	3	4
1	57·14	8·57	8·57	25·71	36·11	16·66	5·56	41·67
2	34·29	14·29	5·71	45·71	27·77	13·89	5·56	52·78
3	59·99	0·00	2·86	37·15	44·44	13·89	8·33	33·33
4	22·86	20·00	5·72	51·43	22·22	2·78	13·89	61·10
5	85·71	2·86	2·86	8·57	75·00	0·00	0·00	25·00
6	88·57	2·81	0·00	8·57	50·00	2·78	5·46	41·67
7	25·71	5·71	2·86	65·70	13·89	11·12	2·78	72·23
8	81·47	3·70	0·00	14·81	66·67	3·70	0·00	29·63
9	87·50	0·00	0·00	12·50	77·77	3·70	0·00	22·23
% mean	57·14	7·50	3·57	31·79	42·36	7·99	5·21	44·44

Table III

Fig. 1, no.	Category 1 % difference (C–E)	Category 4 % difference (E–C)
1	+21·03	+15·96
2	+ 6·52	+ 7·07
3	+15·55	− 3·82
4	+ 0·64	+ 9·67
5	+10·71	+16·43
6	+38·57	+33·10
7	+11·82	+ 6·53
8	+14·80	+14·82
9	+ 9·73	+ 9·73
% mean	+14·78	+12·65

tabulated values representing *differences* between (*a*) the percentage of identifications falling within category 1 for each figure in the E-series, as between the control and the experimental group, and (*b*) the percentage of identifications falling within category 4 for each figure in the E-series, as between the experimental and the control group.

It will be seen from the first column in Table III that the value obtained for every difference is positive. In the second column, the difference is positive in every case but one.

One may conclude that, although the differences between the per-

centages of correct identifications of the E-figures are found to vary considerably among themselves, such differences are nevertheless in every case positive. Similarly, although the values for the differences between the percentages of failures in identification as between the two groups vary considerably from figure to figure, these differences turn out to be positive in every case but one; i.e. in the case of one figure only were fewer failures in identification recorded in the experimental than in the control group. This one negative difference is further extremely small, amounting only to a percentage difference of 3·82.

We may conclude that, although certain figures appear, with regard to subsequent identification, to be more greatly influenced than others by the introduction of an interpolated task, a positive difference in accuracy of identification as between the control and the experimental group occurs in the case of every member of the E-series.

(2) *Certainty in correct identification*

In Table IV are set out the percentage values regarding certainty of identification in the case of correct recognition. The principle according to which degree of certainty has been graded has been explained above.

Table IV

	Grade			
Group	A	B	C	D
C	84·38	6·87	7·50	1·25
E	82·79	7·38	6·56	3·28

It will be seen from the table that there is extremely little difference between the figures obtained from analysis of reports from the two groups. One may conclude that the degree of certainty experienced in the event of correct identification is not reduced by the introduction of conditions designed to favour the operation of retroactive inhibition.

(3) *Immediacy of correct identification*

In Table V are shown the respective percentage values for 'immediate' and 'delayed' correct responses in the case of the two subject-groups.

Table V

Group	Immediate	Delayed
C	76·32	23·68
E	84·07	15·93
Diff.	7·75	

It will be seen that there is a percentage difference of 7·75 between the values for 'immediacy' of correct responses in the two groups: i.e. the percentage of 'immediate' as compared with 'delayed' correct responses is slightly higher in the experimental group as compared with the control group. This result will be discussed below.

(4) *Deviations from accuracy in recognition and recall*

In the above account of the methods used in classifying accuracy of identification, it was pointed out that certain types of deviation in the reproductions occurred with great frequency and regularity.

Such deviations were not, for the purposes of the experiment, held to signify any degree of failure in recognition, and the identifications in question were correspondingly included in category 1. In the majority of such identifications, in fact, recognition was reported as being immediate and certain. This observation nevertheless raises a question, which though incidental as regards the specific problem under consideration, is of considerable interest in relation to a wider object of the investigation, namely to an understanding of the nature of the relationship between recognition and recall. The question raised is the following: Are such deviations mere inaccuracies *in recall*, or can they be regarded as criteria of deviations from accuracy *in recognition itself*? Now the conditions were such that the subjects reproduced, wherever possible, the E-figure in question immediately after termination of exposure of each member of the I-series. It would therefore seem rather unlikely that inaccuracies in the reproductions represent, at any rate wholly, deviations engendered in the actual process of reproducing. It appears by no means improbable, especially in view of earlier findings of the present author, that they represent actual deviations in recognition itself. Some examples of such deviations are given in Fig. 2.

The present writer has also pointed out that recognition may under certain conditions display the same fundamentally constructive character as does recall.[1] Some further evidence in favour of such a conclusion turned up in the results of the present experiment. It has already been mentioned that reproductions were in certain cases found to deviate from the original to a highly significant degree. It was on account of the occurrence of such deviations that it was found necessary to introduce category 2. Some of these cases may now be considered:

(1) One of the I-figures (Fig. 1, no. 1) was so constructed as to contain, in addition to the original E-figure, a figure (which we may call

[1] Compare Bartlett (1) and Zangwill (13).

E′) structurally equivalent to the latter, though oriented differently in space. On the evidence of the reproductions, E′ was 'recognized' as the original by 8·45 % of the total subjects.

(2) Another I-figure (Fig. 1, no. 2) was so constructed as to contain, in addition to the original E-figure, a figure (E′) structurally the mirror-image of the latter. E′ was 'recognized' as the original by 12·68 % of the total subjects.

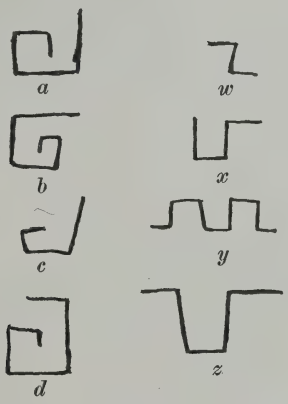


Fig. 2. Four reproductions of Fig. 1, no. 2 (*a*, *b*, *c*, *d*) and of Fig. 1, no. 3 (*w*, *x*, *y*, *z*) demonstrating deviation from accuracy in recall or in identification.

These examples testify to the 'generic', constructive form which recognition may take. It is more than likely that under less rigid experimental conditions and over longer periods of time this more constructive variety constitutes the typical process.

VII. Discussion

It has been shown that recognition of simple figures under the conditions of the experiment is detrimentally affected by the interpolation of conditions known to be conducive to the operation of retroactive inhibition. The first point that must now be considered is the relation between the process of identification as operating under our experimental conditions and recognition as it normally occurs. To recapitulate, our conditions were such that each E-figure was re-presented incorporated in a more comprehensive drawing. There is every reason to believe that, in the absence of specific experimental instructions, no recognition would

have taken place. Nevertheless, when an attitude of 'expectation' and 'search' is induced by previous instruction, we have shown that recognition occurs very frequently, and moreover, that in the majority of cases such recognition follows directly upon exposure of the I-figure. There thus seems good reason to believe that, although we are studying recognition under rather artificial conditions, such conditions are not essentially distorting the nature of the process. It may be recalled in this connexion that Bartlett has pointed out that all recognizing involves the operation of an attitude; probably the attitude factor involved in the present experiment is not very much more specialized than the types of attitude involved in the experiments of Heine and von Restorff. All laboratory experiments of this type are, in fact, bound to involve attitudes of a relatively specialized nature. What is of greater importance is the fact that, owing to rigid standardization of instructions, the attitude factor was the same in the case of both subject-groups. It therefore follows that the difference in results as between the two groups cannot be attributed to a factor of this nature. One may therefore safely seek the cause of this difference in the factor of the interpolated activity, which constituted the only 'variable' in the experiment. From what is known of the conditions of retroactive inhibition, it would therefore appear justifiable to attribute our general result to this cause.

It may be objected that, owing to the nature of the conditions, identification of the E-figures, particularly in cases of 'delayed' identification, might involve such activities as recall, comparison of percept with images, and other higher mental processes. It could be argued that retroactive inhibition might influence the result by preventing the possibility of such recall, i.e. that the inhibition might in reality affect recall only and not directly influence recognition. Some evidence in favour of such a view may be derived from the observation that 'immediacy' of correct responses in the case of the experimental group gave a slightly higher percentage value than in the case of the control group. This difference is, however, insufficient to account for the whole of the difference in accuracy of identification between the two groups. It therefore seems justifiable to assume that if recognition be rendered more discriminative, and consequently more exacting, than is usual, then performance in identification is subject to the detrimental influence of retroactive inhibition.

It must be pointed out that the effect of retroactive inhibition upon identification is rather small as compared with the degree of decrement obtained in experiments designed to assess the influence of this factor

upon recall. In this connexion, von Restorff's emphasis upon the 'specificity' of mental function would appear to be justified. In our experiment, as in Gibson's, identification of the original material constitutes a relatively 'high-grade' accomplishment in comparison with normal recognition. The accomplishment in question may be regarded as a function intermediate in 'specificity' between literal recall and ordinary recognition. Interpolation of a learning activity between presentation and representation appears to influence this function in the same manner and direction as in the case of recall, but to a less pronounced degree. We may tentatively conclude that the amount of retroactive inhibition due to interpolation of a given activity increases with increasing 'specificity' of the memory function under examination.

The results of this experiment justify Koffka's assumption that repetition after an interval will produce some effect on recognition. They also support his interpretation of Dahl's findings as the outcome of retroactive inhibition. Our results do not, however, enable us to decide whether retroactive inhibition is itself due to processes of the nature of 'trace-aggregation'. One may conclude that although our results are in conformity with such an hypothesis, the evidence which they provide in its favour is not by any means crucial.

The main importance of our general result lies in the fact that it provides additional evidence for the view that recall and recognition represent fundamentally similar types of process, influenced in an essentially similar manner by variation in the same types of condition.

VIII. Conclusions

1. If a learning activity be interpolated between exposure of a series of simple figures and their re-exposure as incorporated within members of a series of more comprehensive figures, there is found:

(*a*) A small, though statistically significant, decrease in the percentage of correct identifications;

(*b*) A small, though statistically significant, increase in the percentage of total failures in identification;

(*c*) A slight reduction in the relative number of 'delayed' as compared with 'immediate' identifications.

2. There seems good reason for attributing this detrimental effect of an interpolated task upon identification to the influence of retroactive inhibition. The effect may be due in part to the influence of this factor

upon recall, but there is reason to believe that retroactive inhibition may have a direct effect upon recognition itself. Such a conclusion would accord well with the findings of Gibson and McKinney.

3. No influence of an interpolated task upon the degree of certainty experienced in the event of correct identification can be detected.

4. It is suggested that the extent of retroactive inhibition may vary directly with the 'specificity' of the memory function under investigation.

5. It is pointed out that these results provide indirect evidence in support of the 'trace-aggregation' theory of Koffka and von Restorff and that they support Koffka's interpretation of Dahl's result as a function of retroactive inhibition.

6. An incidental result is the production of further evidence bearing upon the possibility of deviation from accuracy occurring in the actual process of recognition itself.

7. The most important outcome of the investigation is to strengthen the general view that recognition and recall are influenced in an essentially similar manner by variation in the same general types of condition.

My thanks are due to Prof. Bartlett for his encouragement and helpful criticism. I have also to thank Mr E. G. Chambers for some advice concerning statistical treatment of the results.

REFERENCES

(1) Bartlett, F. C. (1932). *Remembering.* Cambridge University Press.
(2) Dahl, A. (1928). "Über den Einfluss des Schlafens auf das Wiedererkennen." *Psychol. Forsch.* xi, 290–301.
(3) Gibson, J. J. (1934). "Retroaction and the method of recognition." *J. gen. Psychol.* x, 234–6.
(4) Gottschaldt, K. (1929). "Über den Einfluss der Erfahrung auf die Wahrnehmung." *Psychol. Forsch.* xii, 1–87.
(5) Heine, R. (1914). "Über Wiedererkennen und rückwirkende Hemmung." *Z. Psychol.* lxviii, 161–237.
(6) Hollingworth, H. L. (1913). "Characteristic differences between recall and recognition." *Amer. J. Psychol.* xxiv, 532–44.
(7) Jenkins, J. G. & Dallenbach, K. (1924). "Obliviscence during sleep and waking." *Amer. J. Psychol.* xxxv, 605–12.
(8) Koffka, K. (1935). *Principles of Gestalt Psychology.* London: Kegan Paul.
(9) Lund, F. H. (1926). "The criteria of confidence." *Amer. J. Psychol.* xxxvii, 372–81.
(10) McKinney, F. (1935). "Retroactive inhibition and recognition memory." *J. exp. Psychol.* xviii, 585–98.

(11) ORMER, E. B. VAN (1933). "Sleep and retention." *Psychol. Bull.* XXX, 415–39.
(12) RESTORFF, H. VON (1933). "Über die Wirkung von Bereichsbildung im Spurenfeld." *Psychol. Forsch.* XVIII, 299–342.
(13) ZANGWILL, O. L. (1937). "An investigation of the relationships between the processes of reproducing and recognising simple figures with special reference to Koffka's trace theory." This *Journal*, XXVII, 250–75.

(*Manuscript received* 4 *March* 1937)

AN INVESTIGATION INTO THE DEGREE OF VARIATION IN THE SCORE OF A MOTOR PERSEVERATION TEST

By JANE DARROCH

I. *Introduction* (pp. 248–250).
II. *Problem* (p. 250).
III. *Subjects* (p. 251).
IV. *Procedure* (pp. 251–252).
V. *Scoring* (pp. 252–253).
VI. *Results* (pp. 253–260).
(*a*) *General description of type of variation of score* (pp. 253–255).
(*b*) *Study of the question whether the variation in score is sufficient to make the test useless as an instrument of measurement* (pp. 256–259).
(*c*) *Comparison of score of first trial with average score for a group of later trials* (pp. 259–260).
(*d*) *Note on the decrease in the score of one subject towards the end of the testing* (p. 260).
VII. *Discussion of result* (pp. 260–261).
VIII. *Summary of conclusions* (pp. 261–262).
References (p. 262).

I. Introduction

The practice of psychologists in the use of motor tests of perseveration has varied in the past in that some have used the result of the first trial, or of early trials, of such a test as a measure of the subject's perseveration, whereas others have discarded the earlier trials and used the average score for a group of later trials. Lankes (6) repeated the letter-writing test, which was the only motor test he used, four times, and used the average score for the last three of these. Bernstein (1), Hargreaves (4), Wynn Jones (9), and Jaspar (5) used that for the first trial; Cattell (2, 3), Stephenson (10, 11) and Yule (12) used that for the first or the first and second. Pinard (7), on the other hand, repeated his tests ten times, discarded the first five trials, and used the average score for the last five, while Rangachar (8) repeated his tests twenty-five times and discarded the first five trials. Pinard's results are invalid in any case (as are Bern-

stein's) because he scored the tests by a method which was almost certainly in part a measure of speed of hand-writing, but it is clear that Rangachar's results are not comparable with those of other investigators.

Pinard's reason for discarding his earlier trials was that the score in these was very variable, and it was still not constant in the trials from the sixth to the tenth, though it varied less. He gives reliability coefficients for the average score for the four tests which he used, the inverted S test, the triangles test, the mirror image test and the alphabet and number test; that obtained by correlating the sixth trial with the seventh was 0·568. A higher one, namely one of 0·724, is certainly given also for the trials from the sixth to the tenth, but this appears to have been obtained by correlating the mean score for two of these trials with that for another two. The reliability coefficients of perseveration tests have usually not been as large as would be desirable. Lankes's very small ones of 0·16, 0·32 and 0·41 for the letter-writing test are of course not strictly reliability coefficients, as different letters were used in the different trials. In Cattell's first published work on perseveration, however, the reliability coefficient of the triangles test is only 0·41, and that of the letter-writing test is 0·52, while he recently obtained reliability coefficients of 0·60 for the *w* test, 0·38 for the triangles test and 0·59 for the *abc* test. Yule obtained coefficients ranging from 0·016 for the *Z* test and 0·330 for the *e* test to 0·522 for the 9 test and 0·523 for the 6 test. The present writer found in an earlier investigation that the score of the *w* test was still very variable by the tenth trial. Rangachar's reliability coefficients are certainly much larger—his lowest is 0·759 for the alphabets test for a group of thirty-five 11-year-old boys, and several are over 0·9—but he does not say how he obtained them, and it seems likely that they were obtained by correlating the average score for ten of the twenty trials, which he retained after the elimination of the practice series, with the average score for the other ten, which would conceal variations in score between individual trials. Stephenson (10) found little change in the scores obtained by mental patients in motor tests when these were repeated after some months, except where the emotional state of the patient had changed too, but results obtained from mental patients may be determined by different factors from those which operate in the case of normal people. All the reliability coefficients mentioned so far are test re-test coefficients; Bernstein obtained fairly high split reliability coefficients, by making his subjects mark each part of each test half-way through the time, but these are irrelevant to a consideration of the variability of the score from trial

to trial. Hargreaves gives the reliability of the reversed stroke test, using the figures 234567, as 0·80, that of the inverted S test as 0·73, and that of the I.T. test as 0·41, but he does not say whether he obtained these coefficients by repeating the tests or by Bernstein's method.

Lack of constancy in the score is a serious fault in any test which purports to measure temperament or character, for these do not vary from day to day, though emotional mood may.

There are two kinds of motor perseveration test, the creative effort test, in which the measure of the subject's perseveration is derived from a comparison of his speed in performing a familiar writing movement with his speed in performing a closely similar but unfamiliar one, and the alternation test, in which the measure of his perseveration is derived from a comparison of the speed at which he can perform two writing movements continuously with the speed at which he can alternate them. Both kinds of test have been used indiscriminately; Cattell (3) claims to have proved that they are measuring the same factor, but his conclusion appears to be invalid. He gave a test which could be scored as either a creative effort test or an alternation test, and obtained a correlation of 0·52 between these two scores for a group of fifty-two subjects. It can, however, be shown that there is bound to be some correlation, for mathematical reasons alone, because of the inequality of difficulty of the two items in the alternating part, even if the subjects experience no more difficulty in alternating them than in writing them continuously.[1] In the present investigation also, a test was used which was scored both as an alternation test and as a creative effort test, because this fact was not realized till after the investigation had been concluded.

II. Problem

The purpose of the present investigation, which was undertaken at the suggestion and under the guidance of Dr Collins, of Edinburgh University, was to find whether the score in a motor test of perseveration became constant when it was repeated a large number of times, or whether it continued to vary. It was decided to give the *w* test on each of fifty more or less successive days to each of a small group of subjects, in order to find whether the score was still variable up to the fiftieth trial. The *w* test was selected because it is a well-known perseveration test used by Stephenson and Cattell, and because it seemed to admit of being scored both as a creative effort test and as an alternation test.

[1] This fact was pointed out to the writer by a student in the Ordinary Psychology Class, Edinburgh University.

III. SUBJECTS

There were sixteen subjects; a larger number did not seem necessary, as the scores of each subject were to be compared only with each other and not with those of other subjects. They were all young adults, ranging in age from 19 to 30. A and B were lecturers in the Psychology Department of Edinburgh University, D is the present writer, E and F were students in the Advanced Psychology Class and G to Q inclusive were students in the Ordinary Psychology Class. Originally, there were two more subjects, C and R, but their results were discarded because they did not do the test under sufficiently regular conditions. A, B, I, J, K, M and P were men and the others were women. G, J, K, L, O and Q had done ten previous trials at the test, P had done five, and D had done a large number, possibly about fifteen, but this did not seem to be a drawback, as it only increased the total amount of practice which these subjects would have had by the end of the experiment.

IV. PROCEDURE

The test was given once on each of about fifty more or less successive days; the number of trials ranged from forty-seven to fifty-one in the case of different subjects. Most of the subjects had an interval of one or two days at week-ends, and an occasional break of three days occurred. It would have been difficult to obtain subjects who were willing to do the test on fifty absolutely successive days, and it was also considered that a test of which the score is affected by the length of time which has elapsed since its last application is too delicate an instrument to be of practical use. The test was done at approximately the same hour at least on the majority of the days. Such changes in time as did have to be made did not actually have any noticeable effect on the score; it did not seem to vary more widely from the average when the hour was unusually early or late, nor was it consistently higher or lower at such hours than at the usual one. The intervals in testing at week-ends did not appear to affect the score of any subject except E, who tended to write ordinary *w*'s more slowly after a week-end and reversed ones faster. Tests were usually timed by the writer, and in those cases in which the subject was timed by someone else the importance of accurate timing was stressed. The writer did not, of course, time her own trials.

The test consisted of three parts separated by rest pauses of 30 sec. In the first or X_1 part, which lasted for 30 sec., the subjects made ordinary *w*'s; in the second or X_2 part, which also lasted 30 sec., they

made reversed *w*'s; and in the third or *Y* part, which lasted 1 min., they made the two sorts alternately.

The subjects were told that they must work as fast as possible, and that the *w*'s must have the two main parts and the hook, and might, but need not, have an extra stroke at the beginning. The writing of both sorts of *w* was demonstrated, and the subjects were also shown a number of sample *w*'s, of which two were, they were told, more or less ideal ones, others just good enough to pass, and others not good enough. This was intended to prevent the subjects from sacrificing quality too much for the sake of speed. The majority of the subjects were told not to count their score, as it was felt that the additional incentive to speed involved in this might cause irregularities; by an oversight E and F were not told this, and E counted hers till the twenty-eighth trial, without the writer noticing that she was doing so. F counted hers at the twenty-fifth and twenty-sixth trials, and was then stopped. A, I, J and K disobeyed the instruction a certain number of times. Perhaps it is unwise to forbid subjects to count their score, as the restriction seems irksome to many, and as it is sometimes possible, as A and J found, to count silently while actually writing the *w*'s. In any case, the scores of those subjects who counted were not consistently more or less variable than the scores of those who did not. A's score was one of the most variable, and J's by far the least. The test was done in most cases on large lined quarto paper; K used a smaller size a few times and F once.

In describing the results it will be necessary to make some mention also of the results of the previous experiment in which some of the subjects took part. In it ten trials of the *w* test were done on four days separated from each other by a week's interval, two or three trials being done on each day; and while Subjects K, L, O, P and Q worked as fast as possible, G and J belonged to a group who were instructed to work at their natural pace. P was absent on the last two days and so only completed five trials. The previous practice done by D was done under irregular conditions, as she was trying the test on herself without any idea that what she was doing would affect the results of any work that was to be published. Her first trial was timed by someone else and the result was kept, but she timed many of the other trials herself and did not record all of these.

V. Scoring

The test was scored by the "ratio" method. Three methods of scoring were used. The first, $\frac{X_1+X_2}{Y}$, was intended to be a pure measure of the

subject's difficulty in alternating the two kinds of w, but it was afterwards found that in a test such as the w test, in which the X_1 and X_2 parts are unequal in difficulty, this score is not really a pure measure of the subject's difficulty in alternation, for it is bound for mathematical reasons to be greater than unity, even when the subject has no more difficulty in alternating the two kinds of w than in writing them continuously. The second method, $\frac{X^1}{X_2}$, was a pure measure of the subject's difficulty in writing the reversed w as compared with his difficulty in writing the ordinary one. When this method is used the test is a pure creative effort test. The third method, $\frac{X_1}{X_2}+\frac{2X_1}{Y}$, which is Stephenson's method of scoring the test (although the present writer's lettering is different from his), was intended as a combined measure of the subject's difficulty in writing the reversed w and of his difficulty in alternating the two w's. Actually, if no difficulty in alternation exists, this score and the score $\frac{X_1+X_2}{Y}$ are both pure but rather complex measures of difficulty in writing the reversed w; if, on the other hand, it does exist, they combine the measurement of it in a certain proportion with the measurement of difficulty in writing the reversed w.

At first an attempt was made to count as an error any w less good than those shown on the sample sheet as sufficiently good to be passed, and subjects were criticized if they made many errors. This, however, was causing Subject I to have a great number of errors and consequently a very irregular score, and it seemed likely that the constant criticism would discourage him, so a more lenient standard was adopted. I, however, still had so many errors that his score became entirely meaningless from the sixteenth trial onwards; his w's were just little wavy or bent lines. Finally the tests were re-scored with all the w's counted as correct except where the subject had repeated the same kind of w twice in succession in the Y part, or had substituted one part of the test for another, as by making alternating w's in the X_2 part. During the latter part of the testing—from about the tenth trial onwards—no subject was criticized at all for making errors.

VI. RESULTS

(*a*) *General description of type of variation of score*

Graphs were drawn of the perseveration scores, by the three methods of scoring, of the sixteen subjects and it was clear from these that the scores varied very considerably. The only scores which became practically

constant were the scores $\frac{X_1}{X_2}$ and $\frac{X_1}{X_2}+\frac{2X_1}{Y}$ in the case of Subject J. Two main kinds of change could be distinguished, namely, gradual changes over a relatively long period, and apparently random day to day fluctuations, either about a constant level or superimposed on the gradual changes. The commonest gradual change was a fairly steep and rapid decrease in the scores $\frac{X_1}{X_2}$ and $\frac{X_1}{X_2}+\frac{2X_1}{Y}$ in the earlier part of the experiment, which was clearly due to the fact that the speed of writing the reversed *w* increased more from practice than the speed of writing the ordinary one. Of the eight subjects who had had no previous practice at the test, four, namely B, F, I and N, showed this to a marked degree, and E to a slight one. J, L, P and Q showed it to a marked degree in the trials done before the present experiment. P, who during the latter part of the present experiment had the lowest perseveration scores of any of the group, was practically unable to make the reversed *w* at all in the five trials, which he did in the previous experiment; he almost always put the hook on the wrong end. D had a high score in her first attempt at the test. The changes in subjects' scores, which took place in the trials done before the present experiment, are not, of course, exactly those which would have taken place in the earlier trials of the present experiment, had the subjects in question done it without previous practice. The results of the previous experiment do, however, in spite of the different conditions under which it was done, contribute to the proof that a pronounced decrease in the scores $\frac{X_1}{X_2}$ and $\frac{X_1}{X_2}+\frac{2X_1}{Y}$ takes place very frequently during the earliest attempts at the test. Nevertheless, this decrease does not occur invariably. The scores of the remaining six subjects either fluctuated about a constant level or showed a variety of different kinds of gradual change. Except with subject H, in whose case the score $\frac{X_1}{X_2}$ fluctuated about a constant level, while the score $\frac{X_1}{X_2}+\frac{2X_1}{Y}$ tended to rise gradually throughout the experiment, the graph of the score $\frac{X_1}{X_2}+\frac{2X_1}{Y}$ was very similar to the graph of the score $\frac{X_1}{X_2}$.

Fig. 1 shows the graph of Subject N's score for the method of scoring $\frac{X_1}{X_2}+\frac{2X_1}{Y}$. It is a typical graph, with a steep initial decrease and a relatively small day to day fluctuation.

Fig. 2 shows Subject M's score for the same method of scoring; it has large day to day fluctuations about an approximately constant level.

For the method of scoring $\frac{X_1+X_2}{Y}$, nine of the subjects, namely A, D, E, F, G, J, K, M and O, had scores which fluctuated about an ap-

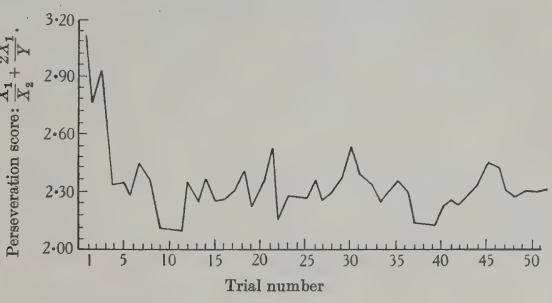


Fig. 1.

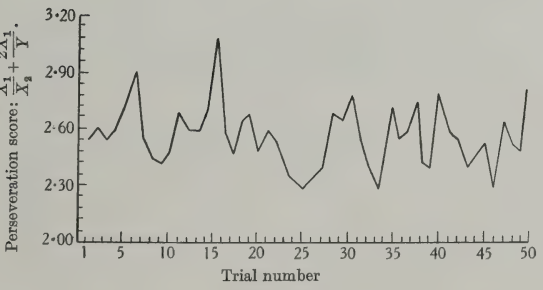


Fig. 2.

proximately constant level throughout the present experiment, and those of G, J, K and O did so throughout the previous experiment also. D had a high score in her first attempt. B's score showed a steep decrease at the beginning of the present experiment, and those of L and Q did so at the beginning of the previous one. I's score showed a gradual decrease throughout the present experiment. H's showed a gradual increase throughout, and those of N and P showed a gradual increase towards the end of the present experiment; in the previous one, as has been already stated, P's score was extremely high for all the methods of scoring.

(*b*) *Study of the question whether the variation in score is sufficient to make the test useless as an instrument of measurement*

The next task was to find whether the variation in the perseveration score was sufficient to make the test useless as a measuring instrument. It was assumed that the true measure of a subject's perseveration was the score which he reached after practice effect had been eliminated, and the first fifteen trials were discarded as a practice series. By this time the initial decrease in the scores $\frac{X_1}{X_2}$ and $\frac{X_1}{X_2}+\frac{2X_1}{Y}$ had ceased except in the case of Subject I; in his case it continued to about the twenty-fifth trial, but it seemed unreasonable to discard more than fifteen trials as a practice series. Next, any of the remaining trials were marked in which anything had occurred, which might possibly disturb the score, such as distraction by noise, fatigue on the part of the subject, or the fact that the test was done at an unusual hour, or after an interval of a day or more on which there was no testing. If the perseveration score for any of the trials so marked was larger or smaller, by any of the three methods of scoring, than any other score for the trials from the sixteenth to the last inclusive, or if the speed of any of the three parts was distinctly lowered in any of the marked trials, the trial in question was discarded. If, however, this was not the case, the trial was retained, and it was assumed that the potentially disturbing influence had not upset the score to more than a very slight extent at most. In other words, the range of the scores of the trials finally retained is determined by trials in which no disturbing influence of any kind was present. The number of trials finally retained ranged from twenty-seven to thirty-four in the case of different subjects.

It was considered that the test would be a sufficiently reliable measure of perseveration, if the variation in score for the trials from the sixteenth to the last was not so great but that the lowest scores obtained by very high perseverators were higher than the highest scores obtained by moderate ones, and the lowest scores obtained by moderate ones were higher than the highest obtained by very low ones. A distribution was accordingly made of the scores of each subject, in order to find whether this was so. Tables I, II and III give the distribution of the subjects' scores for the three methods of scoring. The subjects are arranged in the order of size of their mean perseveration score for the method of scoring in question, which is given.

It is clear that the subjects are not at all well differentiated from each

Table I. *Distribution of subjects' scores for scoring* $\frac{X_1+X_2}{Y}$

Subject...	K	M	H	N	A	B	J	F	G	Q	E	L	D	O	I	P
	Mean score for all trials from the 16th to the last															
	1·08	1·08	1·07	1·03	1·02	1·00	1·00	0·99	0·99	0·98	0·97	0·97	0·96	0·96	0·94	0·93
Class interval:																
1·24–1·26	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.
1·21–1·23	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.
1·18–1·20	1	1	.	.	1	.	.	.	.	.	.	.	.	.	.	.
1·15–1·17	3	1	2	.	0	.	.	.	.	.	.	.	.	.	.	.
1·12–1·14	1	1	7	1	2	.	.	.	.	.	.	.	.	.	.	.
1·09–1·11	12	7	5	4	2	.	.	.	.	.	1	.	.	.	.	.
1·06–1·08	2	11	3	6	5	3	4	.	3	1	1	.	.	.	.	.
1·03–1·05	7	6	4	8	5	14	14	13	3	0	3	1	1	.	.	.
1·00–1·02	4	5	9	6	6	9	12	17	8	12	3	5	6	7	3	3
0·97–0·99	1	.	1	4	3	4	3	2	8	11	5	10	7	8	5	3
0·94–0·96	.	.	1	4	4	.	.	.	4	4	9	7	6	9	4	9
0·91–0·93	.	.	.	.	3	.	.	.	2	4	5	4	4	8	8	10
0·88–0·90	.	.	.	.	.	.	.	.	3	.	2	.	3	.	6	3
0·85–0·87	.	.	.	.	.	.	.	.	.	.	1	.	.	.	1	2
0·82–0·84	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1

Table II. *Distribution of subjects' scores for scoring* $\frac{X_1}{X_2}$

Subject...	K	A	H	L	M	O	Q	F	E	J	N	B	D	G	I	P
	Mean score for all trials from the 16th to the last															
	1·42	1·40	1·38	1·32	1·31	1·27	1·25	1·24	1·23	1·21	1·19	1·15	1·09	1·08	1·03	0·98
Class interval:																
1·65–1·69	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.
1·60–1·64	.	.	0	.	.	.	.	.	.	.	.	.	.	.	.	.
1·55–1·59	3	1	2	.	.	.	.	.	.	.	.	.	.	.	.	.
1·50–1·54	4	4	3	.	.	.	.	.	.	.	.	.	.	.	.	.
1·45–1·49	3	6	3	1	2	.	.	.	2	.	.	.	.	.	.	.
1·40–1·44	5	5	3	0	5	.	.	.	2	.	.	.	.	.	.	.
1·35–1·39	8	7	5	6	3	2	1	3	1	.	.	.	.	.	.	.
1·30–1·34	7	2	7	13	10	8	2	1	2	.	.	.	.	.	.	.
1·25–1·29	1	3	5	5	5	10	13	7	2	1	2	2	.	1	.	.
1·20–1·24	.	2	3	2	5	9	14	13	10	23	12	3	.	1	1	.
1·15–1·19	.	1	.	.	4	3	2	8	7	9	11	8	6	1	4	
1·10–1·14	.	.	.	.	.	.	.	.	1	.	6	14	6	11	2	.
1·05–1·09	.	.	.	.	.	.	.	.	3	.	2	3	9	9	4	2
1·00–1·04	.	.	.	.	.	.	.	.	.	.	.	.	5	6	6	13
0·95–0·99	.	.	.	.	.	.	.	.	.	.	.	.	1	2	5	7
0·90–0·94	.	.	.	.	.	.	.	.	.	.	.	.	.	.	4	6
0·85–0·89	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	2
0·80–0·84	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1

Table III. *Distribution of subjects' scores for scoring* $\frac{X_1}{X_2}+\frac{2X_1}{Y}$

Subject...	K	H	A	M	L	O	Q	F	N	E	J	B	D	G	I	P
	Mean score for all trials from the 16th to the last															
	2·67	2·62	2·59	2·54	2·43	2·34	2·34	2·33	2·31	2·29	2·29	2·22	2·10	2·10	1·99	1·90
Class interval:																
2·90–2·99	1	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.
2·80–2·89	6	4	1	2	.	.	.	.	.	.	.	.	.	.	.	.
2·70–2·79	7	2	10	4	.	.	.	.	.	.	.	.	.	.	.	.
2·60–2·69	7	13	6	4	1	.	.	.	1	.	.	.	.	.	.	.
2·50–2·59	8	4	5	9	5	.	1	1	1	3	.	.	.	.	.	.
2·40–2·49	1	5	4	8	9	9	4	4	4	1	.	1	.	.	.	.
2·30–2·39	1	3	3	6	11	16	16	16	15	6	17	4	.	1	1	.
2·20–2·29	.	.	2	1	1	5	11	11	9	10	16	10	5	4	2	.
2·10–2·19	.	.	.	.	.	2	.	.	4	7	.	15	8	11	5	.
2·00–2·09	.	.	.	.	.	.	.	.	.	2	.	.	10	9	4	5
1·90–1·99	.	.	.	.	.	.	.	.	.	.	.	.	3	6	4	12
1·80–1·89	.	.	.	.	.	.	.	.	.	.	.	.	.	.	8	7
1·70–1·79	.	.	.	.	.	.	.	.	.	.	.	.	.	.	3	6
1·60–1·69	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1

other. For example, when the method of scoring $\frac{X_1+X_2}{Y}$ is used, K and M, who have on the average the highest scores, have each five scores of 1·02 or lower, and P, who has on the average the lowest score, has three scores in the class interval 1·00–1·02. In other words, even after practice effect has been eliminated, a single trial of the *w* test scored by this method does not give a score which can be absolutely relied upon to differentiate a subject, who is on the average an extremely high perseverator, from one who is on the average an extremely low perseverator. For the method of scoring $\frac{X_1}{X_2}$ the situation is slightly better, but it is still far from satisfactory. K, who has on the average the highest score, has one score in the same class interval as one of G's, who is on the average the third lowest perseverator, and twenty-four of his scores fall in the same class intervals as nine of E's, who is the eighth from the bottom. P, the lowest, has two scores in the class interval 1·05–1·09, in which E has three. For the scoring $\frac{X_1}{X_2}+\frac{2X_1}{Y}$, P, the lowest, is again not completely differentiated from E, and K, the highest, is not completely differentiated from I, the second lowest.

It was thought that the test might differentiate the subjects more reliably if the average score for five trials were taken as a measure of the subject's perseveration, and the day to day fluctuations were thus smoothed out. If motor tests of perseveration were ever proved to be a

measure of an important temperamental quality, it would not be an impossibly laborious procedure to give a subject twenty trials at a test, discard the first fifteen as a practice series, and take the average score for the sixteenth to the twentieth as the measure of the subject's perseveration.

The scores were accordingly averaged in groups of five; the trials from the sixteenth to the twentieth, the twenty-first to the twenty-fifth, and so on, were taken together. Owing to the discarding of some trials on account of disturbing influences, five trials could not be included in every average, but averages derived from less than four trials were disregarded in considering the result.

For the method of scoring $\frac{X_1+X_2}{Y}$, K and M were now differentiated from everyone below A, which was satisfactory, but the lowest perseverator had three scores as high as one of A's. For the scoring $\frac{X_1}{X_2}$, P was differentiated from everyone above G, but K was not completely differentiated from anyone above J, and I was not completely differentiated from anyone below Q. For the scoring $\frac{X_1}{X_2}+\frac{2X_1}{Y}$, K was completely differentiated from everyone below M, and P from everyone above D, which was satisfactory; but H, A and M, who are high perseverators on the average, and D, G and I, who are low ones, were still not differentiated from the moderate ones. H's score, for example, now varied from 2·46 to 2·71, that of F, who is a moderate perseverator on the average, from 2·26 to 2·48, and that of I from 1·84 to 2·24. Thus even when the score is an average score for four or five trials, the test does not reliably differentiate high and low perseverators from moderate perseverators.

(c) *Comparison of score of first trial with average score for a group of later trials*

As Lankes, Hargreaves, Wynn Jones, Jaspar, Cattell and Stephenson used the score for the first, or an early, trial as the measure of the subject's perseveration, while Rangachar used the mean score for the trials from the sixth to the twenty-fifth inclusive, it seemed worth while to compare the score for the first trial with the mean score for the trials from the sixth to the twenty-fifth. This could of course only be done in the case of the eight subjects who had had no practice in the test before the present experiment. It was found that not only did the score of these

subjects change considerably in most cases, but their rank order changed considerably too; for example, for the scoring $\frac{X_1}{X_2}+\frac{2X_1}{Y}$, the rank order of the subjects for the first trial was F, B, N, A, E, I, M, H, while for the mean score for the trials from the sixth to the twenty-fifth it was A, F, M, H, B, E, N, I. While a result thus derived from eight subjects is suggestive rather than conclusive, the fact that the rank order changes so much suggests that the perseveration scores obtained by Rangachar, on the one hand, and, on the other, those obtained by investigators using the score for the first, or an early, trial, are not both measuring perseveration equally well.

(*d*) *Note on the decrease in the score of one subject towards the end of the testing*

A fact which casts further doubt on the power of the test to discriminate reliably high perseverators from lower ones is that Subject A's score began to decrease pronouncedly towards the end of the period of testing. The three lowest of his scores shown in the distribution in Table I, and the six lowest shown in Tables II and III, were all obtained after the thirty-ninth trial. It seems probable that if, say, seventy trials of the test had been done instead of fifty, he would have appeared as on the average a moderate perseverator, and the scores of single trials might even have represented him as a low one; his score $\frac{X_1}{X_2}+\frac{2X_1}{Y}$ for the forty-seventh trial is 2·21. Certain facts suggest that the decrease in score was due to his discovering an easier way of making the reversed *w*; from about the fortieth trial he began to show a tendency to substitute curves for angles in its formation. At the end of the forty-seventh trial he gave the introspection that he found the reversed *w* easier, and seemed to have learnt something new, and that he thought he was making the hook and the rest of the reversed *w* in one movement. It seems as if those subjects who have been making high perseveration scores, that is, those who find the reversed *w* difficult to write, are liable at any moment to show a pronounced decrease in score because they have found an easier way of writing the reversed *w*.

VII. Discussion of result

The present investigation was carried out on the basis of the assumption that the true measure of a subject's perseveration is the score which he obtains after practice effect has been eliminated, and it has shown that

this score is too variable for the test to be a reliable measure of any quality. A case can, however, be made out for the view that the true measure of a subject's perseveration is the score which he obtains at his first attempt at a perseveration test. If those workers be excepted who have used the "difference" method of scoring without eliminating the factor of speed of writing, there are, in all the published work on perseveration, only twelve correlations between perseveration tests that are more than four times their probable error, and these are all between first or second attempts at tests. Eight were obtained by Cattell (3), three by Wynn Jones (9) and one by Hargreaves (4). The correlation of 0·52, obtained by Cattell between the reversed stroke test scored as an alternating test and the same test scored as a creative effort test, is omitted as spuriously high, because the score of the former would not be a pure measure of the subject's difficulty in alternation, owing to the inequality in difficulty of the two items. It is possible that the correlations between motor tests of perseveration are caused by some such factor as initial adaptability to making an unfamiliar movement or series of movements, which is measured only by the earliest attempts at a test. Cattell (2) indeed suggests that this is the case; he says that "Pinard has maintained that valid perseveration measurements can only be obtained by giving the subjects practice and throwing away the practice results", but that "there seems to be nothing in the intrinsic nature of perseveration requiring such conditions" and that "one would expect the best measurements to be obtained when the test is entirely unfamiliar".

VIII. Summary of conclusions

1. When sixteen subjects did approximately fifty trials of the *w* test on fifty more or less successive days, their perseveration score varied so much that high and low perseverators could not be reliably discriminated from moderate ones, either by means of the scores for single trials, or by the average scores for groups of four or five trials.

2. In the case of nine of the subjects there was, for two out of the three methods of scoring used, a pronounced decrease in score in the course of their earliest attempts at the test, which was clearly due to the fact that the speed of writing the reversed *w* was more affected by practice than the speed of writing the ordinary one.

3. The rank order of eight subjects, who had never had any previous practice at the test, differed considerably when the score for the first

trial was compared with the mean score for the trials from the sixth to the twenty-fifth.

4. The score of one subject, who was on the average a high perseverator, decreased markedly after the fortieth trial.

REFERENCES

(1) Bernstein, E. (1924). "Quickness and intelligence." *Brit. J. Psychol. Monograph Supplements*, III, No. 7.

(2) Cattell, R. B. (1933). "Temperament tests. I. Temperament." *Brit. J. Psychol.* XXIII, 308. "II. Tests." *Brit. J. Psychol.* XXIV, 20.

(3) Cattell, R. B. (1935). "On the measurement of perseveration." *Brit. J. educ. Psychol.* V, 76.

(4) Hargreaves, H. L. (1927). "The 'faculty' of imagination: an enquiry concerning the existence of a general 'faculty' or group factor of imagination." *Brit. J. Psychol. Monograph Supplements*, III, No. 10.

(5) Jaspar, H. H. (1931). "Is perseveration a functional unit participating in all behaviour processes?" *J. soc. Psychol.* II, 28.

(6) Lankes, W. (1915). "Perseveration." *Brit. J. Psychol.* VII, 388.

(7) Pinard, J. W. (1932). "Tests of perseveration. I. Their relation to character." *Brit. J. Psychol.* XXIII, 5. "II. Their relation to certain psychopathic states and to introversion." *Brit. J. Psychol.* XXXIII, 114.

(8) Rangachar, C. (1932). "Differences in perseveration among Jewish and English boys." *Brit. J. educ. Psychol.* II, 199.

(9) Spearman, C. (1927). *The Abilities of Man*, Chap. XVII. London: Macmillan. (Gives results obtained by Wynn Jones and not published elsewhere.)

(10) Stephenson, W. (1931–32). "Studies in experimental psychiatry. I. A case of general inertia." *J. ment. Sci.* LXXVII, 723. "II. Some contact of *p*-factor with psychiatry." *J. ment. Sci.* LXXVIII, 315. "III. *P*-score and inhibition for high-*p* praecox cases." *J. ment. Sci.* LXXVIII, 908.

(11) Stephenson, W. (1934). "An introduction to so-called motor *P*-tests." *Brit. J. educ. Psychol.* IV, 186.

(12) Yule, E. P. (1935). "The resemblance of twins with regard to perseveration." *J. ment. Sci.* LXXXI, 189.

(*Manuscript received* 21 *January* 1937)

A PRELIMINARY INVESTIGATION OF THE INTERVIEW

By S. M. HARVEY

From the Cambridge Psychological Laboratory

I. *Introductory analysis of problems approached* (pp. 263–264).
II. *Preliminary experiments* (pp. 264–265).
III. *Character bias experiment* (pp. 266–284).
(1) *Method and type of bias* (pp. 266–270).
(2) *Quantitative treatment of results* (pp. 270–275).
(*a*) *Relation of ratings to the record sheets presented* (pp. 270–273).
(*b*) *Individual variability of interviewers in regard to the effects discussed* (pp. 273–275).
(3) *Qualitative treatment of results* (pp. 275–284).
(*a*) *Consistency of the behaviour of candidates and interviewers* (p. 275).
(*b*) *Discussion of results in relation to interviewers' comments on ratings* (pp. 275–279).
(*c*) *Interviewers' comments on the experiment* (pp. 279–280).
(*d*) *Tendencies which appear to have operated in the assessments of interviewers* (pp. 280–284).
IV. *Summary* (pp. 284–286).
Appendix (pp. 286–287).

I. Introductory analysis of problems approached

In so far as we respond to people and form opinions about them, personality and character diagnosis is part of our everyday life. We make pragmatic judgments which, though largely unformulated, enable us to some extent to predict the behaviour of others in those situations in which we have to deal with them. Whatever the processes that underlie these judgments, it seems justifiable to assume that they are the same in kind, though possibly more formalized and exact, when such diagnosis has to be made in an interview. Hence the study of the interview may provide valuable data for an approach to the wider problems of the psychological basis of such diagnosis and its validity.

It is possible to make a rough differentiation between interviews which are 'personally related', referring specifically to a candidate's character and constitution, and those which are 'impersonally related', e.g. interviewing for testimony—though it must be remembered that personal prestige of the witness or candidate is always likely to be to some extent involved, even in the second type, and that there can be no sharp dividing line between the two. In brief, both types involve a perceptual situation, complicated by the social relationship between interviewer and candidate.

The impersonally related interview has been widely studied (see Moore's summary of work on this subject[1]). This work has been primarily applied to factors which affect the witness, but it is obvious that the interviewer will himself be affected by similar factors. It is not the purpose of this paper to discuss the investigations quoted; but among their results the question of attitude—whether arising from cultural bias or from more individual tendencies to interpret immediate events in terms of personal experience—stands out from these, as well as from more theoretical work on remembering, as vitally important, and requiring further investigation. If character is socially determined, then we may infer that certain aspects of expression—movement, voice production, and attitude as shown in verbal behaviour—will not only be promoted and emphasized in members of a given social group, but their interpretation will be socially determined. Hence, irrespective of the symptomatic reliability of such expression (this is a separate problem) we may expect to find the operation of biases due to personal experience and social group determination not only in interviewer and candidate in the impersonally related interview, but also in the interviewer in the personally related type. Such biases or attitudes must operate in the pragmatic judgments of which we have spoken, and probably also the manner of their formulation. When we know the conditions of the operation of pre-established bias, and its effect upon an interviewer's judgment, we shall be in a better position to examine the reliability of cues used in character and personality diagnosis and the whole question of their interpretation. For this reason, the question of 'bias' or pre-established attitudes in an interview was selected for study in the investigations reported here.

II. Preliminary experiments

It was hoped that a preliminary investigation of the manner of operation of a bias in an impersonally related interview, which is in many ways simpler to approach, might yield information valuable in tackling the

[1] Moore, E. H. (1935). "Elements of Error in Testimony." *J. app. Psychol.* XIX, 447.

personally related type. But initial attempts to set up or to emphasize experimentally a bias in regard to the interpretation by the interviewer of the information to be obtained from his candidate yielded little new information which could be directly used in the approach to the personally related interview. The difficulty encountered in biasing an interviewer, as compared with the ease with which the candidate can be biased in regard to simple or even more complex perceptual material, was in itself a fact of some significance. In one experiment upon the effects of a pre-established social group bias (political), the selective, stereotyping effects of such bias upon the reconstruction of the relevant material was confirmed; and some indication was obtained that agreement or disagreement of bias between interviewer and candidate also affected the reconstruction. There were observable differences between the results when a right-wing interviewer examined a left-wing candidate (or vice versa) about a political speech, and when their views were in accord. The effect operated mainly through the difference in the social relationship of interviewer and candidate, and yielded indication of the importance of the social relationship involved in an interview, and a possible line for its investigation. In another experiment the effect of a bias that the candidate 'might be lying' was examined; and it was found that in some instances such a bias did have a marked effect upon the interviewer's judgment as to the candidate's veracity. Behaviour which in a different context might merely have been regarded as due to shyness was, in conjunction with the warning contained in the instructions, interpreted as evidence that the candidate was lying. On the other hand, if there was no marked shyness and the interviewer's general impression of the candidate was good, the bias might be rejected, even though the candidate was in fact lying. That is to say, the effect of the bias depended upon the relation of the interviewer's pre-established ideas of the behaviour to be expected of people who are being truthful or untruthful, to the behaviour shown by the candidate.

It is probable that more information on this subject could be obtained from a study of the effects of 'bias' in very simple perceptual situations. Certainly much more experiment on these lines is needed. But in the meantime, the results served to indicate that it is possible to introduce a bias experimentally in a personally related type of interview, and it was upon this that the main experiment was undertaken.

III. Character bias experiment

The aim of this experiment was, while impressing upon the interviewer the importance of his personal opinion, to examine the effect upon his judgment of his candidate's character of a favourable or unfavourable impression about the candidate given beforehand by means of a prepared 'record sheet'.

(1) *Method and type of bias*

Interviewers. Five young post-graduates (code names I.1 to I.5), including one young married woman, took part. All had had some experience of dealing with boys in social work.

Candidates. Six boys were obtained through the local Juvenile Employment Committee. None had left school more than a year. Some were still seeking work, others were already placed. The authority was so kind as to provide information about the boys' characters, derived from school and other relevant history. Unfortunately it was not possible to obtain a group with a regular distribution of character quality, bad to good. Brief records of the boys who served (code names C.a. to C.f.) are to be seen in Table I, p. 271.

Material. Five short cartoon stories were used, taken from comic papers. In each the hero indulged in some kind of anti-social behaviour, and a wrongdoer "got away with it".

Instructions. The boys were told only that it was an experiment on giving evidence, and that they were to read slowly through the story twice. After this they were taken to the interviewer, required to make a report, and then left.

Interviewers were given before the interviews started an 'Introduction to the Experiment' (meant in part to deflect attention from its real purpose). In this it was described as an attempt to test out a new, indirect method of approach to character diagnosis. Insufficiency of time in interviewing for vocational guidance, except by special institutions, was emphasized, and the quickness with which this method might be applied. The futility was pointed out of direct questioning on points likely to reveal character quality, since answers would "be likely to reflect not the impulses the boy would obey in practical life, but his conscious opinions, or those of which he supposes his interviewer would approve". They were informed of the nature of the comic strips, and told that it was hoped that they would, by observing the boys' behaviour generally, their attitudes to interviewers, task and stories, etc., be able

to assess them in regard to the enumerated character traits. No detailed suggestions were made as to method. In final instructions, presented before each interview, these requirements were summarized, and they were told that the record sheets, which were simultaneously presented, were intended to help them in choosing the best questions to ask, etc. They were informed that these were provided because questioning the boy about himself was not allowed. All interviewers asked whether they were to refer to the Records when making their ratings, and were told that their *independent* ratings were required. They were allowed to ask some questions of the boys about their attitudes to 'critical' points in the stories. There was no rigid time limit, but interviewers were asked to try to get their information if possible in the space of 15–20 minutes.

Quality ratings required. Interviewers were given after each interview a rating sheet in the following form. (Authorities who undertake vocational diagnosis have to estimate boys for these qualities, and it was thought that the untrained interviewer would also be able to make such ratings.)

1. General Reliability:
   (*a*) steadiness and absence of distractibility;
   (*b*) stability and solidity of character;
   (*c*) suggestibility or submissiveness to influence of people, ideas or example.
   Summary of reliability:
2. General Sociability:
   (*a*) ease in getting on with people;
   (*b*) aggressiveness;
   (*c*) lack of ease in social situations.
   Summary of sociability:
3. Emotional Stability:
   (*a*) cheerfulness and alertness;
   (*b*) excitability;
   (*c*) dullness or slowness in response.
   Summary of emotional stability:

Ratings were asked in the following grades:

A: boys who possess a quality (good or bad) to an outstanding degree—*very marked*;
B: boys whom you would say were *above the average* in respect of any quality, but in whom it is not shown to the marked degree of the A class;
C: boys who show a quality to an *average* or moderate extent;
D: boys who show a quality to a *small extent*;
E: boys who show a quality *very little*, or to an imperceptible extent.

Interviewers were asked to give on a separate sheet comments on the basis of, or the manner in which they arrived at, their conclusions and ratings.

Biases and standards of comparison. The reports obtained from the authority already referred to were used as the basis of information given in the record sheets presented to the interviewers. That relating to the three qualities to be rated was embodied under two or more heads. The scheme was as follows: the record presented before the first interview for each boy and interviewer was to be in accord with this report, a 'standard' on all points. In the remaining four records for each boy, his standing in regard to these qualities was varied in two stages or in two directions. Each quality was intended to be biased twice only. Reliability and sociability were to be varied independently; when one was biased, the other to be maintained at the standard level. Emotional stability was to be varied once in conjunction with reliability (both towards the average or upward), and once with sociability (both towards the average or downward).

Thus, e.g. a boy whose 'standard' reliability was *high* appeared once with an *average* record, and once with a *low* record, for that quality. If his standard reliability were *average*, he appeared once with a *high*, and once with a *low* record for that quality. The other two qualities were treated similarly.

Sample record sheets are given in the Appendix (see p. 286), to show the manner in which the relevant information was introduced and changed. Provided the standards were reliable, the detection of the effect of the bias was, under this scheme, a simple matter of comparison. For even if the standard itself acted as a bias, deviations due to the other biases would be obvious. Unfortunately, as it turned out, the original standards were not reliable in all cases.

The reports used in compiling the standards derived indirectly from the reports of the heads of the elementary schools which the boys had attended. These included marks for: regularity, punctuality, character, conduct; sociability, self-confidence, aggressiveness, leadership, and general stability; school subjects and woodwork. Notes were provided as to home circumstances, posts held, and other relevant data. Comments used in the record sheets and ratings used as standards of comparison had to be derived from this information. Unfortunately, marks were not all given by the same schoolmaster, and in the course of contact with the boys the experimenter came to the opinion that the reports had been in some cases misinterpreted in relation to each other. As a result it was found necessary to adopt new standards of comparison in regard to certain qualities for some of the boys. In arriving at these, reference was made to the ratings derived from the authority's statements, to ratings

given independently by the experimenter[1] (it cannot be denied that though she had fuller information she was probably subject to bias similar in kind to that to which interviewers were subjected!); and to an average of the ratings given by all interviewers, primary importance being given to the unbiased, but consideration also being given to the biased, occasions. The record sheets were maintained in their original form, as the experiment was already under way. This means that certain boys were never interviewed after presentation of a completely unbiased record sheet, and that the carefully planned distribution of biases was much altered.

Control. Unfortunately, owing to the changes discussed above, and to the fact that one of the boys originally engaged fell out before any interview but after times for the whole group had been fixed, and another after his third interview, so that two new boys had to be introduced to the scheme, the original scheme for control of the position factor in regard to the order in which boys were interviewed, presentation of biases, etc., broke down, and control is regrettably poor.

Comparison of ratings and records. The records were, of course, framed in verbal terms. Because of the difficulty of formulation of fine shades of difference, only three main categories were used and implied in the information related to each quality: very good or high; good or average; fair or low. Five possible categories were provided for interviewers, in the hope that compromise between real and biased quality would show up better. In practice the extremes were avoided, and it seems probable that 'very good' in the record must be compared with anything above average (C) given by an interviewer, and so on. For this reason in the ensuing comparisons A to B, and C to D, are grouped together and ratings derived from the records similarly treated.

Sufficient data (standards) for the full investigation of the effect of the biases on the subsidiary scores required of the interviewers was not available, but it was hoped that by inviting this further articulation of

[1] In her own assessment of the boys, the experimenter attempted to obtain an impression of their quality in the following respects:

Reliability: related to the question "would I trust him?" and embodying such factors as conscientiousness, steadiness and independence of thought and behaviour;

Sociability: "how friendly is he?", inferred from such things as the boy's responsiveness to friendly remarks, more positive qualities such as advances on his own initiative, or his shyness or reserve;

Emotional stability: "how would he react in an emergency?"—derived chiefly from an impression of the boy's degree of responsive animation, dullness or impassivity, quality of voice, tenseness in the strange experimental situation, and such behaviour as being 'muddled' in replies at the beginning of any interview, or showing nervous movements.

ratings a better idea would be available of the manner in which the bias operated, if at all. These subsidiary ratings are disregarded in the following discussions.

In Table I, p. 271, are shown the order in which the boys were seen by each interviewer, the record presented with each boy, and the ratings given by interviewers.

(2) *Quantitative treatment of results*

(*a*) *Relation of ratings to the record sheets presented.*

In all, 28 interviews were conducted and recorded stenographically. If we regard each record as covering 3 items, 84 records were presented, of which 42 were standards and 42 biased in some direction. Ratings agreed with the record sheets in a total of 43, i.e. about 50 % of cases. Agreement was distributed as follows:

	Records presented		Agreement with records	
	Standard	Biased	Standard	Biased
Reliability	13	15	12	11
Sociability	15	13	8	2
Stability	14	14	6	4

That is to say, about 60 % of agreement was shown with standard, and about 40 % with biased, records. The crude results given above might suggest equally: a difference in the ease with which judgments on the three qualities are made; or that sociability was least amenable to bias and reliability most so; or that the bias given in regard to reliability was stronger; or finally that some extraneous factor (*vide* low agreement with the standards) was affecting both sociability and stability ratings.

Examination of the relation of interviewers' ratings to the strength of bias given yielded the information that proportional agreement was as follows:

Favourable biases:

+2 (DE raised to AB): one case only, agreement resisted.

+1 (C raised to AB, or DE to C): 11 records presented, agreement 45 %.

Unfavourable biases:

−1 (AB reduced to C, or C to DE): 22 records presented, agreement 27 %.

−2 (AB reduced to DE): 8 records presented, agreement 75 %.

Table I. *Order in which candidates were interviewed, records presented in each case (standards in plain type, biases italicized), with difference of bias from the standard (in brackets), and rating given by interviewer (alphabetical)*

Interviewer and quality		(1)		(2)		(3)		(4)		(5)		(6)	
		C.a.		C.b.		C.c.		C.e.		C.f.			
I.1	Reliability	High	B	High	B	*Low* (−1)	C−	*Average* (−1)	C+	*Average* (−1)	B+	—	
	Sociability	*Average* (−1)	A−	*Low* (−2)	B	Average	C−	Average	C−	*High* (+1)	B+	—	
	Stability	*Average* (−1)	B+	*Low* (−2)	C	Average	D	*High* (+1)	C−	Average	A	—	
		C.b.		C.d.		C.e.		C.a.		C.c.		C.f.	
I.2	Reliability	High	B	*Average* (+1)	C−	High	A	*Low* (−2)	D	*High* (+1)	C	High	B
	Sociability	High	B	Average	D	*High* (+1)	C	*Average* (−1)	B	*Low* (−1)	C	Average	B
	Stability	High	C	*High* (+2)	D	Average	B	*Average* (−1)	B	*Low* (−1)	C	*High* (+1)	B
		C.d.		C.e.		C.a.		C.c.		C.b.		C.f.	
I.3	Reliability	Low	B	*Low* (−2)	C	High	A	*High* (+1)	A	*Average* (−1)	B	*Average* (−1)	C
	Sociability	*Low* (−1)	B	Average	A	High	A	Average	B	High	A	*Low* (−1)	C
	Stability	*Average* (+1)	B	Average	B	*Average* (−1)	A	Average	B	*Average* (−1)	B	*Low* (−1)	B
		C.d.		C.e.		C.a.		C.b.		C.c.		C.f.	
I.4	Reliability	Low	D	High	B	*Average* (−1)	C	*Low* (−2)	D	*High* (+1)	C	*Average* (−1)	C
	Sociability	Average	C	*Low* (−1)	D	*Average* (−1)	B	High	D	*High* (+1)	D	Average	C
	Stability	Low	D	*Low* (−1)	B	High	C	High	C	Average	C	Average	C
		C.e.		C.a.		C.b.		C.c.		C.f.			
I.5	Reliability	High	B	High	B	High	B	Average	C	*Low* (−2)	C	—	
	Sociability	Average	C	*Low* (−2)	B	*Average* (−1)	B	Average	D	Average	B	—	
	Stability	Average	C	*Low* (−2)	C	High	B	*High* (+1)	C	Average	C	—	

Standards (as amended for comparison) for the six candidates

	C.a.	C.b.	C.c.	C.d.	C.e.	C.f.
Reliability	High	High	Average*	Low	High	High*
Sociability	High*	High	Average	Average	Average	Average
Stability	High*	High	Average	Low	Average	Average

* Amended.

Thus, as would be expected, the stronger negative bias gives the highest proportion of adoption. Unfortunately only one case of double positive bias was possible. But of the 'equal' favourable and unfavourable biases, the former obtained the greater measure of agreement. Further enquiry showed that of the 6 cases of agreement with the strongest unfavourable bias, 4 ratings were a compromise between standard and bias, and that full agreement (−2) was shown only in regard to reliability. Finally, whereas reliability obtained the greatest measure of acceptance of unfavourable bias, adoption of the favourable biases was more equally distributed between the three qualities.

Table II. *Number of records presented (standard and biased) as compared with the number of ratings given by interviewers in each grade and for each quality*

	Records presented	Agreement with records		Ratings given	Ratings expressed as percentage of records presented
		Total	%		
A–B (high):					
Reliability	13	11	84	14	107
Sociability	7	4	57	15	214
Stability	8	2	25	13	162
Total	28	17	60	42	150
C (average):					
Reliability	8	6	75	9	112
Sociability	15	5	33	8	53
Stability	14	5	36	12	85
Total	37	16	43	29	78
D–E (low):					
Reliability	7	6	85	5	71
Sociability	6	1	17	5	83
Stability	6	3	50	3	50
Total	19	10	53	13	68

The results summarized in Table II are somewhat conflicting. Aggregate agreement of ratings with the records presented is highest for the A–B records (60 %), which suggests a tendency towards favourable ratings, but agreement with the D–E records is fairly close (53 %). The groupings of ratings (see columns 4 and 5) about the upper end of the scale supports the suggestion of such a favourable tendency. On the other hand, the totals obscure marked differences in degree of agreement shown with the records for the three qualities. Moreover, if we regard the standard, unbiased, ratings of the boys as the 'expected' ratings, then such a grouping would be expected as their records were on the whole above rather than below average. But (see Table III *b*) the 'expected' number of A–B ratings would be 40, of C ratings 38, and of

D–E ratings 6. That is to say, the number of average ratings given was markedly less, of A–B ratings slightly more and of D–E ratings markedly more, than would be expected (see column 4). The implications of these conflicting tendencies do not become clear until we turn to the examination of the results of individual interviewers.

(*b*) *Individual variability of interviewers in regard to the effects discussed.*

In Table III the number of records presented to each interviewer in the three categories is compared with the number of ratings given. It will

Table III

(*a*) *Records presented to and ratings given by individual interviewers in the three grades*

	Reliability		Sociability		Stability		Totals	
	Records	Ratings	Records	Ratings	Records	Ratings	Records	Ratings
I.1								
AB	2	3	1	3	1	2	4	8
C	2	2	3	2	3	3	8	7
DE	1	.	1	.	1	.	3	.
							15	15
I.2								
AB	4	3	2	3	3	3	9	9
C	1	3	3	2	2	2	6	7
DE	1	.	1	1	1	1	3	2
							18	18
I.3								
AB	2	4	2	5	.	6	4	15
C	2	2	2	1	5	.	9	3
DE	2	.	2	.	1	.	5	.
							18	18
I.4								
AB	2	1	2	1	2	1	6	3
C	2	3	3	2	2	4	7	9
DE	2	2	1	3	2	1	5	6
							18	18
I.5								
AB	3	3	0	3	2	1	5	7
C	1	2	4	1	2	4	7	7
DE	1	.	1	1	1	.	3	1
							15	15

(*b*) *Total standard ratings for all boys in the three grades*

	Reliability	Sociability	Stability	Total
AB	4	2	2	8
C	1	3 (or 4)	3	7 (or 8)
DE	(1)	.	(1)	. (2)

(Brackets indicate numbers for standards of records presented to 3 interviewers who saw 6 boys.)

be seen that I.1 and I.3 do show a comparative grouping of ratings at the upper end of the scale; that this tendency is slighter in I.2 and I.5 who, at any rate in the gross scores, maintain almost the original groupings; while I.4 shows a marked tendency towards either the average or the lower end of the scale. The facts might be taken to indicate either that when the ratings were made there were in operation tendencies particular to the individuals concerned, or alternatively that an impression which determined the rating given on one quality for a given boy spread also to affect the ratings given on the other two qualities. Both suggestions will be supported in the later qualitative analysis. In the meantime, reference back to Table I does reveal uniformity in the three ratings in a number of cases which seems sometimes to derive from the standard quality of the candidate, and sometimes from the reliability bias, and which suggests the importance of a general impression which spread in its effects over the three ratings.

Table IV. *Comparison of agreement shown by individual interviewers with standard and with biased records (favourable and unfavourable)*

	Reliability			Sociability			Stability			Totals		
	Agree	Disagree		Agree	Disagree		Agree	Disagree		Agree	Disagree	
Interviewer		+	−		+	−		+	−		+	−
I.1												
Standards	2	.	.	2	.	.	.	1	1	4	1	1
Biases:												
Favourable	.	.	.	1	.	.	.	.	1	1	.	1
Unfavourable	2	1	.	.	2	.	1	1	.	3	4	.
I.2												
Standards	3	.	.	1	1	1	.	1	1	4	2	2
Biases:												
Favourable	1	.	1	.	.	1	1	.	1	2	.	3
Unfavourable	1	.	.	.	2	.	.	2	.	1	4	.
I.3												
Standards	1	1	.	2	2	.	.	2	.	3	5	.
Biases:												
Favourable	1	.	.	.	.	.	1	.	.	2	.	.
Unfavourable	2	1	.	.	2	.	.	3	.	2	6	.
I.4												
Standards	2	.	.	2	.	1	3	.	2	7	.	3
Biases:												
Favourable	.	.	1	.	.	1	.	.	.	.	.	2
Unfavourable	3	.	.	1	1	.	.	1	.	4	2	.
I.5												
Standards	4	.	.	1	1	1	3	.	.	8	1	1
Biases:												
Favourable	.	.	.	.	.	.	.	.	1	.	.	1
Unfavourable	1	.	.	.	2	.	1	.	.	2	2	.
	23	3	2	10	13	5	10	11	7	43	27	14

To some extent the grouping of ratings might be due to the fact that the number of favourable and unfavourable records presented were, unfortunately, not the same for all interviewers. This, however, could not be the only determinant when the number of records in each grade given to I.3 and I.4 were so closely similar, while that of their ratings is so entirely opposed. Moreover, the direction of agreement and disagreement shown by different interviewers with favourable and with unfavourable biases supports the thesis of differences in individual tendencies in assessment (see Table IV).

(3) *Qualitative treatment of results*

(*a*) *Consistency of the behaviour of candidates and interviewers.*

In order to know whether differences in ratings could be accounted for by differences in the behaviour of boys on separate occasions, careful shorthand notes were made on the behaviour of both interviewers and candidates. *Interviewers* varied from each other, but were on the whole consistent throughout the experiment. Their attitudes are to some extent reflected in the individual tendencies already discussed. I.1, I.2 and I.3 were extremely friendly, making positive attempts (smiles and remarks) to facilitate ease in their candidates. I.5 was responsively friendly, but less positively so. I.4 was moderately friendly, but more formal in his remarks and treatment of the boys than any other interviewer. *Boys* of average sociability and stability were particularly affected by the differences between the interviewers' attitudes, showing a greater or lesser degree of nervousness. Except in this respect, though they varied considerably from each other, they were fairly consistent in their behaviour over the five interviews. The difference in the social relationship may, of course, have been one factor contributing to the greater variability of the ratings in regard to sociability.

(*b*) *Discussion of results in relation to interviewers' comments on ratings.*

Sample schedules follow for three boys: C.a. who could be described as 'good all round', C.c. 'average all round' and C.d. the 'difficult case'. These were chosen because of the differences between the boys, but the comments of the interviewers are typical of the remarks made by each upon any boy.

C.a. This boy was good-looking, open-faced, pleasant-spoken but independent in his comments. He was completely at ease throughout the experiment, and positively friendly with all interviewers except I.5 with whom he was nevertheless responsively friendly. His glance was uniformly open and direct, he was cheerful in manner and alert to answer and give information, animated but self-controlled. There were no

obvious signs of nervousness or excitability, though with I.4 and I.5 he made slight restless movements at the outset of the interview. His account was good and coherent in every case, though his style was vivid and colloquial. Answers to questions were immediate and thorough. In regard to 'critical questions' (those about the moral implications of the story and his attitude to them) his attitude was matter of fact, though he showed slight embarrassment in regard to these with I.1 (his first interview) and I.5. Attitude in recounting critical aspects of stories was matter of fact; he showed some amusement at these with I.2, I.3 and I.5. With I.4 he expressed sympathy with critical characters but gave intelligent reasons. With I.5 he expressed disapproval independently of critical actions, and with I.2 and I.3 agreed in disapproving.

Reliability. Biases presented to I.2 (low) and I.4 (average)—both adopted.

Comments: I.1. "Few significant changes of expression." Not concerned about the man's dishonesty, but what boy of his type would be?" "Stuck to most of his points."

I.2. "Found him charming." Rating a "mere guess. Seemed to evade discussion of ethical points" (though as with I.1 boy dealt with these in a matter of fact way). "Seemed very able to absorb suggestions" (his answers were in fact accurate).

I.3. "Most of my conclusions were arrived at rather from the boy's general attitude than in particular reference to his story. Appearances are notoriously misleading on occasions, but, perhaps with the notes about him in mind, these were exceptionally strong ones." "Went through story quickly...remarkable amount of detail. Thorough and keen to get it right." "Although he quite patently considered the story to be impossible...clear he thought pretty poorly of those who got money for nothing." "Almost told me how the interview would go...quite nicely" (referring to resistance to suggestion).

I.4. "Story was rather less coherent than the 2 boys already seen." "Normal sort of response to ethical judgments, neither accepting my judgments unless he agreed nor showing...contra-suggestion."

I.5. "Did not marshall his facts quickly...gesticulated thoroughly." "Seemed very sure of his own opinions."

Sociability. Biases presented to I.1, I.2, I.4 (all average), and I.5 (low)—but none was adopted.

Comments: I.1. "Clearly interested in the job in hand and co-operated willingly."

I.2. "Delicious twinkle when we shook hands and made me feel 'entirely at home'. Seemed to find roguish delight in attacks perpetrated in the story." "Very good initiative in telling stories."

I.3. "His general attitude." "A little more (aggressive) than a normal boy but not unpleasantly so...also his ease in getting on with people would naturally make him seem and possibly feel a little cock-sure at his period of development."

I.4. "More at ease in telling his story (than 2 previous boys)."

I.5. "Was quite 'matey' and did not choose his words carefully...willing to enter into argument or conversation. His slight egotism may have been caused by nervousness, for he was obviously nervous at the beginning...and I think his alacrity to depart showed that he was still nervous at the end."

Stability. Biases presented to I.1, I.2, I.3 (average)—all resisted, and I.5 (low) adopted in compromise rating.

Comments: I.1. "Showed no signs of nervousness, was cheerful and very alert, perhaps a little excited in the unusual environment."

I.2. "Seemed very cheerful indeed, very quick perception." (Excitability) "should perhaps be B, from the way in which he seemed to understand the stories, he saw the fun in them, though did not overestimate the amount." "Very quick response to all questions."

I.3. "Remarkably alert and pleasant. Very cheerful but not...hearty about it. Excitability seemed perfectly normal."

I.4. "Seemed to be very cheerful, more so than the previous two, with no exceptional excitability or slowness...of response."

I.5. No comment.

C.c. A nice-looking boy, but with a squint. Reserved and shy in manner, and somewhat over-deferent (his frequent use of 'sir' was noted by all interviewers). Very responsive with I.3 but with other interviewers only moderately so. Showed considerable signs of nervousness and a negative excitability: tended throughout to avoid direct exchange of glances; his voice was monotonous and in the intervals of questions he tended to lapse into impassivity which suggested boredom, though when answering he was alert and cheerful. With I.5 showed slight restless movements of the face. His account was in general average good and coherent (only moderately so with I.2) and he showed some evidence of initiative in regard to the task. Answers were to I.3 and I.4 immediate and thorough, to I.2 slower but careful, to I.1 and I.5 brief. With I.2 (2nd interview) he showed some embarrassment at critical questions, but was otherwise matter of fact in reserved way. With all interviewers except I.2 he showed amusement at critical aspects of stories; with I.2, I.3, I.4 and I.5 he agreed in disapproval of critical actions, and with I.1 showed indications of a realization of moral implications of stories but made no special comment. With I.2 and I.5 he expressed sympathy with the critical characters, but gave no explanation.

Reliability. Biases presented to I.1 (low) adopted, and I.2 and I.4 (high) resisted and I.3 (high) adopted.

Comments: I.1. "Oh the influence of the record sheet! Every nervous look is construed as evidence of shiftiness." "He needed a lot of prompting but was usually moderately certain when a suggestion of mine was wrong."

I.2. "Mainly guess. Way of telling the story was quite good." "Not very shocked at immorality of story, but fairly stable ideas in general." "Yielded to opinions given, but probably because it was impolite to contradict."

I.3. "I think the most satisfactory and pleasant boy of the 4 I have so far interviewed." "In classing him A I am afraid intuition has played more part than reason."

"Went very quickly but without scamping it...seemed determined to get them off before he forgot anything." "Showed signs of having an accurate idea of justice (quoted detail of attitude to critical aspects deduced from boy's remarks)." (Suggestibility) "Normal. Reacts...almost noticeably to the attitude of others."

I.4. No comments. A general impression is indicated by comments in regard to sociability and stability.

I.5. No comments.

Sociability. Biases presented to I.2 (low) resisted, and I.4 (high) resisted.

Comments: I.1. "I am surprised that he is as easy to get on with as he appears (from record)."

I.2. "Rather shy." "Seemed rather unhappy in his chair." "Yes, sir." "Quite good in summarizing the positions and characters of the persons in the story."

I.3. "In spite of shyness 'B' because his shyness is of the type that attracts rather than discomforts other people....A very easy manner after the first few minutes." "Slightly less than normally aggressive. It seemed, one had a feeling too slight to justify, that he would react quickly to provocation from others."

I.4. "The most significant thing about this boy seemed to me to be his addressing me as 'sir' repeatedly, suggesting a certain submissiveness and lack of ease."

I.5. No comment.

Stability. Biases presented to I.2 (low) resisted, and I.5 (high) resisted.

Comments: I.1. "Two bright boys previously perhaps tend to make me rate him low."

I.2. "Some sense of humour in comment on the story. By no means exuberant. Excitability not noticeable." "Answered questions instantly and correctly. Gave comment sometimes when asked for it."

I.3. "Both cheerful and alert to a pleasing degree. Excitability normal." "Very responsive, cf. instinctive response to the attitude of others."

I.4. "He was easily the most phlegmatic of the boys, although not really dull with it."

I.5. No comment.

C.d. A nice-looking boy, very pleasant-spoken. One of the two 'problem' cases originally engaged to balance the group. Only completed 3 interviews and was then transferred to a special home away from this district. Was quiet and shy throughout. Very responsive to I.2 but only moderately so to I.3 and I.4. Showed marked signs of nervousness and excitability: tended uniformly to avoid direct gaze from I.2, and some avoidance with I.3 and I.4. Was with I.2 very impassive. Marked restless movements of the whole body and face with all 3 interviewers. Tended to be impassive except when answering and making account, when he was very alert and cheerful. His whole behaviour supports the diagnosis by the local clinic of emotional difficulties. His account was very good (he is highly intelligent) and coherent, with evidence of initiative in regard to the task. Answers to questions were with I.4 slow but careful, with I.2 and I.3 monosyllabic. With all three interviewers he showed marked embarrassment at the critical questions, and refused to discuss the critical aspects of stories. With I.3 he showed sympathy with the critical characters, but refused to make any comment. His behaviour suggested the embarrassment that might be expected in regard to discussion of critical aspects from a boy who had been, as he was then, on probation for theft.

Reliability. Bias presented to I.2 (average) adopted.

Comments: I.2. "He stuck to what he had said." "Consistent in the narrative." "Resists 'petty temptation to lie' but otherwise not judgeable by me." (Note that this referred to stability of character formation, and indicates a remarkable penetration.)

I.3. "Putting him B all through, I should be ready to trust him, except on his evaluation of justice, that would seem to be a fault out of which he could easily be educated." "Kept quite steadily to the point, but not to an A extent" (indicating resistance to suggestion). "Refusal to admit it curious to find a solitary girl hiker."

I.4. "He accepted all the suggestions I made with the exception of the occasion when I asked the first about his reaction to the story" (when the boy showed slight embarrassment which was not apparently noted). "Steadiness (rated B) was indicated by the easy way he went on with the story."

Sociability. Bias presented to I.3 (low) resisted.

Comments: I.2. "Did not look at me." "Nervous twitching of mouth, and hand movements." "Willing to stop his account when I made the slightest sign of wanting to say something (taken to indicate submissiveness)."

I.3. "A rather diffident, shy manner. Definitely very uncomfortable at first." "Showed slight signs of aggressiveness on being asked too much for his opinions on the stories."

I.4. "Sociability I found hard to judge. He seemed nervous without being bashful. I think he would be at home amongst a group of his own social class."

Stability. Biases presented to I.2 (high) resisted, and I.3 (average) adopted.

Comments: I.2. "Stopped talking unless pressed on." "No signs of excitability." "Did not attempt to answer any questions which did not concern the actual picture story."

I.3. "Definitely quite cheerful, especially over things that interested him most." "Good smiling eyes." "Not very excitable, though he did light up a bit...." "General attitude."

I.4. "His response seemed to me to be quick because he was intelligent, but his continued playing with his hands seemed to suggest excitability and a certain degree of emotional instability."

(*c*) *Interviewers' comments on the experiment.*

Standardized questions were asked of all interviewers after their last interviews. The first asked for comment on the experiment and whether they thought this 'picture-method' might be useful in short interviews. Replies all indicated that the experiment had been taken seriously, though most thought that the same information could be obtained in other ways, possibly more easily. The second question enquired whether they had found difficulty in making ratings, and whether they considered their criteria had remained consistent throughout the interviews. All noted difficulty in making the ratings. I.1 noted the effect of previous interviews on later ratings. I.2 "came to judgments about the boys... but difficult to fit the judgments into categories. Of course I really could not tell things like reliability, excitability, aggressiveness, and only *assumed these from general judgments of their character*". I.3 thought difficulty arose "chiefly because I tried to judge boys by appearance and manner more than answers". I.4 noted that "you so rarely come across any one to whom you wish to give the two limits, and feel that you need more than the ratings B and D" (*sic*) and that difficulty was "partly due to lack of sufficient data to make all the judgments". I.5 also noted this

difficulty. I.4 was also not certain what was meant by each character detail. All except I.1 thought their criteria had been consistent.

The third question was as follows:

"It has been pointed out that the record sheets, which were given you (*vide* instructions quoted)... did in fact result in a biased attitude towards the boys. Do you feel that this is true? What was your own attitude towards the record sheets, and the information contained therein?"

On the whole interviewers did not seem to think that they had been much biased by the record sheets. I.1 replied that it "helped to a certain extent in the attitude I took up.... Don't think it biased ratings, except when I found my own impression was very different, in which case I would think twice before putting the rating which I decided on". I.2 thought "on the whole I relied on my own judgments. I didn't think much about the record sheet at all" (this does, in fact, seem likely from the relative independence of most of his ratings). I.3 noted "At first I definitely referred to the character thing afterwards. There may have been a tendency to consider the boy bright or aggressive and so on through having read it first, but later I paid less attention to the record than with the first one or two." (Note that he does show effects of bias in the final as well as the second interview.) I.4 said "When I was inclined to depart from it, I made a much more severe checking of the grounds for departing from it than when I tended to agree with it." I.5 is specially interesting: "I felt that... it might give a biased attitude and tried hard to forget or overlook it completely. It seemed to me that it would be better to regard every boy as average... that gave me an unbiased attitude. Except with the last boy, who was below average in reliability, I completely ignored the record." (Note that all I.5's ratings, except for C.f., concur with the record.) "Temptation to use the record came when I didn't know what to put... but then I tried again not to use that but to find my own valuation." Reference to the ratings given does certainly support this independence.

On the whole, therefore, these comments suggest that, except in the case of I.5, the biases operated without awareness of the interviewers that they were so influenced, or even despite resistance to them.

(*d*) *Tendencies which appear to have operated in the assessments of interviewers.*

(1) *The importance of the general impression obtained of a candidate.* It has already been noted that there appears (*vide* Table I) a certain similarity or uniformity in the ratings given any boy on the three separate

qualities, which suggested either a general impression which influenced all ratings, or a spread to other qualities of an impression gained in regard to one of them (particularly, of course, from reliability). Interviewers' comments as to the basis of their ratings amply confirm this supposition. Apart from the comments on the three cases quoted, for instance, I.1 made specific reference to such an impression in three cases when a reliability or a sociability bias was resisted. I.3's comments were long, though they do not seem to contain any more explicit reference to detail than those of other interviewers, and they abound in evidence of his dependence upon such a general impression—see, for instance, his comments as quoted, or that in regard to C.f. (adopting unfavourable reliability bias) "a very ordinary boy". I.4 made few direct references to such an impression, but his comments suggest it. I.2 made such references, as e.g. of C.e. "general impression very stable, but no particular reason for assuming that", though his comments on the whole group were more carefully articulated, and he referred to a wider range of behaviour detail, than any other interviewer.

Even the nature of comments on particular behaviour detail support this contention of the importance of the general impression. Those upon reliability were more highly articulated than those upon sociability or stability. It is probable that the nature of the experimental material and the instructions made articulation of the cues more obvious. Reference was easy to such detail as manner and consistency of account and answers, attitude to critical aspects of the stories and to the task. It is, however, not the actual detail quoted but its general configuration that is important in relation to the rating given and to the bias. It seems possible, though it is not proved, that such comments were more often explanations or rationalizations *ad hoc* than the real basis of diagnosis.

Comments in regard to sociability are very general indeed. Only I.2 made any consistent attempt to articulate the behaviour cues to which he referred—as in his use of boys' expressed attitudes to the social relationships given in the stories. It appears that, though the validity of generalizing about a boy's general sociability on the basis of his behaviour in an experimental situation might be questioned and though his relation with the interviewer may affect that behaviour, sociability is a quality of function that can be appreciated and even rated, but that at the same time articulation of the things which distinguish a very sociable boy from a less friendly one is difficult. It is of its nature a 'general impression'.

For the untrained interviewer the conception of emotional stability, on the other hand, seems difficult to grasp and to assess. Comments do

little more than repeat the wording of the scheme given for the ratings: "seemed very cheerful", "excitability normal". Such relevant cues as nervous movement were often missed; cues related to stability and sociability were often confused in their application; and intelligence was often confused with alertness. I.4 (who, with I.1, had had some psychological training) gave throughout the most articulated comments in regard to stability.

(2) *Relation of the bias to the 'general impression'.* It does appear from the material that the bias, by evoking an attitude towards the candidate, determined the behaviour to be expected and the interpretation to be given to that observed. Note, for instance, I.1's comments on C.c. on all points, or the condoning, explanatory attitude adopted by I.3 in regard to C.d. Even when whatever behaviour might be expected as a result of the biased attitude was not directly observed, the interpretation of that which was found was affected. Compare the notes of I.1 and I.2 in regard to the reliability of C.a., and the similarity of the comments of I.4 and I.5 as compared with the difference of ratings given C.a. When the biased attitude was directly challenged by a candidate's behaviour, it was either rejected or appeared to be isolated in its effect to a single rating—that which had been specifically biased in the record sheet in most cases, as seen in the comments on C.a. of I.2 and I.4, or on C.d. by I.3, who isolated an unfavourable rating under a subsidiary head.

(3) *Differences in the effects of biases in regard to the three qualities.* At the beginning of Section III (2), differences were noted in the crude results in regard to the degree of agreement with records, standard or biased, for the three qualities. Since then one of the possible explanations advanced, that some extraneous factor was affecting the sociability and stability ratings, has been confirmed. In a number of instances the apparent independence with which these assessments were made was accounted for by the spread of a general impression deriving from reliability bias. Again, the facts:

(*a*) that whereas acceptance of the strongest unfavourable bias was highest, acceptance of the single unfavourable bias (i.e. one move from the standard) was lower than acceptance of a single favourable bias (though this fact must be in part due to the general favourable tendency in assessment noted);

(*b*) that whereas of the biases in regard to reliability more unfavourable than favourable ones were adopted, treatment of biases in regard to the other qualities was more even—acceptance being about equal in either direction;

suggest that the strength of bias applied was an important determining factor. In fact, records contained 'instances' of the ways in which boys had been shown to be unreliable or unsociable, etc., and unwittingly that in regard to extreme unreliability was so chosen as to carry a greater weight of social disapproval than such instances as were chosen in regard to sociability or stability. This was not so for 'average' quality. So unfortunately the observed difference must in part be attributed to a fault in the experimental set-up.

The facts remain, however, that (*a*) in a number of instances sociability resisted 'spread' and reliability bias was isolated in its effect; and (*b*) sociability bias when applied alone tended to be resisted; whence we may conclude that there was real independence in assessment of sociability. (Owing to the special difficulties affecting the results in regard to emotional stability, already discussed, this quality will not be dealt with again.) It seems that there may, therefore, have been a real difference in the ease with which reliability and sociability were assessed, and in their respective amenability to bias. This may perhaps be due to a difference in the way in which the two assessments have to be made. Ordinary experience would suggest that sociability is a function which can be observed in the immediate situation, and experienced *directly* by an interviewer in his relation with the candidate. Reliability, on the other hand, must be *inferred* from a complex pattern: the way the candidate meets the interviewer's eyes, expression, the whole bearing: a pattern which is hard to define, but nevertheless definitely experienced. Both 'patterns' must be assessed by reference to subjective norms, but sociability has in some sense a more definite objective configuration. If this hypothesis is correct, we might expect that the more nebulous pattern would be more amenable to bias, which might create a conflict in the mind of the interviewer in regard to its interpretation. In this experiment interviewers had available more objectively checked cues also, such as manner and consistency of account, resistance to suggestion, attitude to task; and the greater articulation of comments in regard to reliability consisted largely in the inclusion of these cues. But acceptance of bias carried with it the weight given to these cues in the interpretative 'configuration'. The whole question of the relation of objectively checked cues to those which are referred to more subjective standards is raised and requires further investigation.

This interpretation of the results cannot account for the fact that sociability and stability were particularly affected by 'spread'; but here it is probable that we must refer to the intimacy of the relation between

sociability, if not stability, and the 'general impression' that a candidate gives.

(4) *Differences in the tendencies observed in the results of individual interviewers.* Reference to the comments shows that these bear out the rough differences observed in the quantitative results of the several interviewers. The grouping of ratings towards the ends of the scale may now be seen to be in some part due to the operation of a spread of general good or bad impression. I.1 and I.3, however, tend to be favourable in their comments, though this is more marked in the latter's. I.4's comments are not very revealing, though it will have been noted that his more formal, critical, attitude and method in dealing with the boys is in conformity with the tendency observed to group his ratings towards the average or unfavourable end of the rating scale. I.2 gave, on the whole, the most independent ratings, and his comments were, as already noted, more articulated than those of other interviewers. It is interesting that of all the interviewers he had had least experience in dealing directly with boys of this age. Even when adopting bias (*vide* his comments on C.a.) he seemed sensitive to the real (standard) quality of the boy. This suggests the speculation that differences between interviewers is a matter of temperament, though much more investigation is required upon the point.

IV. Summary

The size of the group studied, and the inadequacy of experimental control, preclude any but tentative conclusions being drawn from the results of these investigations. We set out to examine the effect of a bias introduced into instructions given to an interviewer upon his attitude to his candidate, and the effect of this attitude upon his diagnosis of certain character traits in his candidate. We found that such a bias did (without awareness of the interviewer) affect judgment in about 40 % of cases in which it was applied. On the other hand, when true records were presented, there was agreement with these in approximately 60 % of cases. This difference suggests that there is some fundamental basis of reference to behaviour and appearance in character diagnosis which sets a limit to the operation of the biased attitude. This hypothesis is supported by the fact that the interviewers showed a greater degree of independence in their assessments of sociability than of reliability. Assessment of emotional stability seems to have presented too difficult a task for the untrained interviewer—probably because it is unrelated to the normal requirements of his dealings with people.

We noted the importance of a general impression gained of the candidate, which was sometimes such as to result in resistance to bias, and sometimes the mechanism through which the bias appeared to operate. Such general impressions, often called 'halo' effects, have already been noted to affect the diagnosis of personal qualities (see particularly in this connexion Magson[1]). But it seems possible from the results of the present investigation that such an impression is a fundamental condition of diagnosis. It is particularly interesting when considered in relation to the differing degree of articulation on the part of interviewers as to the basis of their judgments between sociability and reliability. Though ratings on the former trait were more independent, articulated comment was less.

We have assumed that such a bias as that used in this experiment operates through its effect upon the attitude of the interviewer. There seems evidence in the results that this is so, and that this attitude results in a search for relevant cues, selection among those found, and what can best be described as special organizations of the cues observed. Differences in the comments on the same boys between occasions when standard and when biased records were presented before the interview appear to lie less in the particular details of behaviour noted, than in the manner of their presentation, and the emphasis given to the same behavioural cues. The significance of the same cues may, as was observed also in the preliminary experiments, change with the context in which they are observed.

In interviewing for character diagnosis at the present time, the assumption is made that there are certain constellations of behavioural cues which have a diagnostic significance, and that interviewers may be trained in the observation of these cues. The results of this experiment suggest that a fuller knowledge is needed of the circumstances in which articulation of such cues is mere rationalization, or the emergence of dominant components in a much wider scheme which may be affected by temporary or extraneous conditions, and in what circumstances they are likely to be consistently reliable. Until this is known, the value of practice or training and the significance of inter-reliability between groups of interviewers similarly trained, is not clear.

It is impossible to dissociate the effects observed in this experiment from those of suggestion; but the indications given of a limit to the operation of a bias do suggest that to some extent other fundamental processes have been involved and can be studied by this method. Hence it seems

[1] Magson, E. H. (1926). "How we judge intelligence." *Brit. J. Psychol., Mon. Suppl.* No. 9.

possible that it is one which may prove fruitful in the approach to the way in which the mental organizations which underlie character and personality diagnosis are established, and in which they become differentiated and formulated.

I would like to record my gratitude to Prof. Bartlett for very much helpful criticism. I wish also to express my appreciation of the co-operation of the subjects, and of the assistance given by the Cambridge Juvenile Employment Committee in obtaining candidates.

APPENDIX. SAMPLE RECORD SHEETS

Candidate b

1. *Standard throughout.*

Schoolmaster's Report (Elementary School):

Character: very good. The boy is exceptionally conscientious, honest and truthful.
General Stability: High. Bright and responsive; thoroughly stable and reliable.
Sociability: High. Very popular with schoolfellows and generally friendly.
Self-confidence: High. Thoroughly self-reliant though not opinionated or stubborn.
Aggressiveness: Average. Comes forward under provocation, but is too friendly and even-tempered for aggressiveness to be marked.
Leadership: High. In general takes an equal part with his fellows, but has a generally responsible attitude.
Conduct: Very good. Sometimes given to high-spirited larking; but generally very responsible and thoroughly reliable. School Prefect. School attendance regular and punctual.

School subjects and woodwork: Fairly good to good. Industrious.

Home Circumstances: Very good. Financially secure and standard high. Parents take great interest in the boy and are very proud of him.

Now in first post.

2. *Reliability reduced to 'Low'.*

Schoolmaster's Report (Elementary School):

Character: Fair. Not very conscientious. Inclined to be untruthful and gives way too easily to temptation.
General Stability: High. Bright and responsive.
Sociability: High. Very popular with schoolfellows and generally friendly.
Self-confidence: Average. Self-reliant but not stubborn.
Aggressiveness: Average. Comes forward under provocation, but is too friendly for aggressiveness to be marked.
Leadership: Low. Takes a fairly equal part with other boys in school activities, but is much too easily influenced.
Conduct: Fair. Though not always reliable and trustworthy (is in fact on probation for theft), will probably improve under supervision.

School subjects and woodwork: Fairly good to good. Needs to concentrate.
Home Circumstances: Good. Financially secure and standard normal. Parents take an interest in the boy and are worried about his unreliability.
Now in second post.

3. *Sociability and emotional stability reduced to 'Low'.*

Schoolmaster's Report (Elementary School):
Character: Very good. The boy is exceptionally conscientious, honest and truthful.
General Stability: Low. Is easily over-excited, sometimes moody. Some signs of nervous strain.
Sociability: Low. Very unequal in his attitude. Not very popular.
Self-confidence: High. Very self-reliant, though tractable.
Aggressiveness: High. Even-tempered on the whole but inclined to respond too quickly to provocation from his schoolfellows.
Leadership: Average. Takes a fairly equal part with his schoolfellows in group activity; but inclines to be solitary.
Conduct: Very good. Somewhat uneven in his attitude and behaviour, but very reliable on the whole. School attendance regular and punctual.
School subjects and woodwork: Fairly good to good. Uneven.
Home Circumstances: Very good. Financially secure and standard high. Parents take a great interest in the boy and are very proud of him.
Now in first post.

Note. It is impossible to give more than three sample record sheets; but the "instances" chosen to illustrate any quality, and the general structure of information, were the same for all boys at any level. With minor necessary alterations, the records of "good" and "bad" boys were interchanged when constructing the biases.

(*Manuscript received* 14 *September* 1937)

PROLEGOMENA TO A PSYCHOLOGY OF ART[1]

By W. R. D. FAIRBAIRN

I. *The artist* (pp. 288–290).
II. *Types of activity* (pp. 290–291).
III. *Art as a play activity* (pp. 291–292).
IV. *The puritan attitude to art* (pp. 292–293).
V. *Art and the dream* (pp. 293–295).
VI. *Inspiration* (pp. 295–299).
VII. *Symbolism* (pp. 299–300).
VIII. *Art-work and the principle of restitution* (pp. 300–303).

I. The artist

The term 'Art' is ordinarily employed to describe a social phenomenon embracing three component elements: (1) The *Work of Art*, (2) The creative *Artist*,[2] (3) The *Percipient* or Audience. In addition to these three factors there is, of course, the fourth factor of *Technique*, which provides the means whereby the artist is enabled to embody his conceptions in actual works of art. A complete psychology of art would thus have to be based upon the study of a fourfold field comprising (1) the artist, (2) the percipient, (3) the work of art, and (4) technique. The psychological aspects of technique must, however, be regarded as falling outside the province of the general psychologist, except in so far as they are determined by the emotional make-up of the artist. Consequently the general psychologist is left with the artist, the work of art and the percipient as the material for his study. Whilst a complete psychology of art would involve bringing each of these three elements under review, it is a matter of importance to decide which of them ought to be made the starting-point for investigation; for it is always possible that one of them

[1] A paper read before the Scottish Branch of the British Psychological Society on 6 March 1937.

[2] In what follows, the word 'artist' will be understood not in the narrow sense of 'painter' (so familiar in common parlance), but in the general sense of 'artistic creator' (irrespective of the department of art in which artistic creation manifests itself). It is true that the thesis of the present paper has been illustrated almost exclusively by examples drawn from pictorial art; but this fact is determined solely by considerations of expository convenience. It must in no sense be regarded as restricting the application of the argument to any one department of art.

may provide a clue, which the others do not, to the real significance of art as a psychological phenomenon. If this be indeed the case, we may at once dismiss the work of art, since the only possible concern of the psychologist with the work of art arises out of its capacity to express the purposes of the artist and to produce effects in the mind of the percipient. The psychologist is thus left with a choice between the artist and the percipient in deciding where to seek a foundation for the psychology of art.

In the past the choice of the psychologist has usually fallen upon the percipient—with the result that hitherto the psychology of art has been largely a psychology of art-appreciation, i.e. a psychology of 'aesthetics'. That this should have been the case is not difficult to understand in view of the historical preoccupation of the general psychologist, at first with purely introspective methods, and later with those of the laboratory; for it is obvious that art-appreciation lends itself to the application of such methods in a way in which artistic creation certainly does not. This concentration of psychological interest upon the percipient has, however, been attended by such a remarkable barrenness of results that one is inevitably driven to the conclusion that it has led the psychology of art into a cul-de-sac. That it should have done so might have been predicted from the first; for obviously it represents an attempt to base the psychology of art upon a relatively unessential element in the total phenomenon concerned. Assuming that the shipwrecked Robinson Crusoe was an artist, it would be ridiculous to maintain that there was no art in the Juan Fernandez island until the arrival of the man Friday. Whereas art is unthinkable in the absence of either the artist or the work of art, there is at least no theoretical difficulty in conceiving of art in the absence of an audience.

Once we reach the conclusion that no secure foundation for a psychology of art is to be found in the study of the percipient, the issue becomes simple. We are left with no alternative but to approach the psychology of art by way of the psychology of the artist—a path of approach, it will be noticed, that has always been taken by adherents of the psychoanalytical school of thought. In his recent book entitled *Art and Society*, Herbert Read makes the statement that "art begins as a solitary activity"; let us amend this statement by saying that art begins as a solitary activity *on the part of the artist*. It then becomes our first task in the present enquiry to determine the general nature of artistic activity. In performing this task, it will be noted, we shall be killing two birds with one stone; for, since artistic activity can be defined only with reference

to the work of art, the work of art must obviously be the complement of artistic activity; and it follows that, in determining the nature of artistic activity, we shall also be determining the nature of the work of art.

II. Types of activity

Activities may be regarded as falling into two great classes: (1) activities undertaken for their own sake, i.e. activities undertaken for the satisfaction provided by the activities themselves; (2) activities undertaken for ulterior ends, i.e. activities undertaken as a means of providing satisfactions independent of those inherent in the activities in question. Another way of drawing the distinction would be to say that, whereas activities of the first class are undertaken 'for fun', those belonging to the second class are 'serious' activities. Or again, employing psycho-analytical terms, we may describe activities of the first class as predominantly determined by 'the pleasure principle', and those of the second class as predominantly determined by 'the reality principle'. It would also seem that we should be justified in regarding the distinction between the two classes of activity as corresponding (roughly, at any rate) to that between 'play' and 'work'. Games as such provide an obvious example of the first class of activity; for we 'play' games for 'fun' at the instance of 'the pleasure principle'. On the other hand, in so far as occupational activities are undertaken with the object of earning a living, we must regard them as belonging to the second class; for they are 'serious' activities determined by 'the reality principle'; and they fall under the category of 'work'.

Any given form of activity may, of course, fall now into the one class, now into the other. Thus, whereas people hunt foxes for fun in England, in the Hudson Bay country fox-hunting is a serious business. Furthermore, any given activity may fall into both classes at once in varying proportions; i.e. it may fall at any point on a scale, at one end of which lies the 100 % 'play' activity, and at the other end of which lies the 100 % 'work' activity. Indeed, most activities would seem to fall at an intermediate point on this scale. So far as artistic activity is concerned, we must regard it as falling into the 'play' class in proportion as the artist seeks the pleasures of artistic creation; but it must be regarded as falling into the 'work' class in proportion as the artist harnesses his art to some ulterior motive such as making a living, gaining social prestige, or propagating a cause. To judge by the majority of opinion recently expressed by art-critics, the portrait-painter Winterhalter would appear to have been more influenced by ulterior motives and would, therefore, fall nearer

the 'work' end of the scale than such portrait-painters as Titian or Velasquez, whose pictures bear the impress of the pure joy of artistic creation. It may perhaps appear anomalous at first sight to apply the term 'play' to the 'work' of the latter artists. Nevertheless, pure artistic activity, which is free from all ulterior motives, is seen to fall automatically into the class of play-activities under the present classification.

III. Art as a play activity

That artistic activity is essentially a play-activity could hardly be better illustrated perhaps than in G. K. Chesterton's essay entitled 'The White Horses'. In this essay, Chesterton describes how he spent a day motoring, with a silent man of peasant stock as his chauffeur, round the traditional battlefields of Alfred the Great. In the course of his pilgrimage he passed three of the numerous white horses carved out by prehistoric men on the chalk hill-sides of the Southern Downs; and he tells us how the sight set him wondering, "Why did barbarians take so much trouble to make a horse nearly as big as a hamlet; a horse who could bear no hunter, who could drag no load?" The answer came later. "As I rolled away out of that country", he continues, "I was still cloudily considering how ordinary men ever came to want to make such strange chalk horses, when my chauffeur startled me by speaking for the first time for nearly two hours. He suddenly let go one of the handles and pointed at a gross green bulk of down that happened to swell above us. 'That would be a good place', he said." Chesterton's chauffeur had entered into the spirit of the prehistoric artist. He wanted to carve white horses on the hillside for fun.

Once we come to regard artistic activity as falling essentially within the class of play-activities, it is not a difficult task to discover what differentiates it from other members of its class; for artistic activity is not just 'doing something'; it is also '*making* something'. *Artistic activity* may, therefore, be defined forthwith; it consists simply in *making something for fun*. On the basis of this definition, we are also in a position to define 'a work of art' and 'art-appreciation'. A *work of art* is something that is made for fun; and, conversely, *anything that is made for fun* must be regarded as a work of art. *Art-appreciation* consists in *perceiving something that has been made for fun*. And furthermore, *aesthetic pleasure* may be defined as *the fun of perceiving something that has been made for fun*.

The conception of art implied in these definitions is doubtless unfamiliar. It certainly involves an enormous enlargement of the sense in which the term 'art' is conventionally accepted. In accordance with this

conception, for example, we must regard as works of art the 'bird's nest' effects produced by the child of three years of age in his first attempts to use a pencil for the fun of making something.[1] It also follows that the sandpies which the child makes on the beach and the ordered arrangements of stones and tins which he makes in the garden are no less genuinely works of art than the sculptures of Michael Angelo.[2]

In accordance with our definition, a work of art may also be an object appealing to any of the various sense departments. Conventionally the term 'work of art' is used in a sense which restricts its application almost exclusively to objects appealing to the visual and auditory senses. The reasons for this restriction cannot be considered in the present paper. Suffice it to say that it is part of the general cultural process whereby the senses of sight and hearing are exalted at the expense of senses, like taste and smell, which are more closely related to the libidinal gratifications of childhood, and, further, that this cultural process is intimately related to the process of repression in the individual mind. In spite of conventional restrictions, however, it must be recognized that, if anything made for fun is a work of art, such an object does not cease to be a work of art because its appeal is to the gustatory or olfactory sense, rather than to the sense of sight or hearing. Consequently there is no justification for denying that aesthetic pleasure enters into the experience of the gourmet, or for withholding the title of 'artist' from the *chef* who derives satisfaction from making his *soufflés* dainty, ephemeral as these works of art may be.

IV. The puritan attitude to art

However unfamiliar the conception of art now suggested, its substantial accuracy is borne out by a consideration of the puritan attitude towards art. The essence of puritanism lies in an extreme intolerance of all forms of activity inspired by pleasure-seeking motives. Consequently the puritan is an arch-enemy of fun; like the melancholic, he is so obsessed with guilt (or, to use psycho-analytical terms, so much under the domination of a tyrannical super-ego) that, in his eyes, all fun is equivalent to sin. If he is not, like the melancholic, quite incapable of being happy, he feels, at any rate, that to be happy is to be wicked. The fact is that he recognizes no activity as legitimate unless it is undertaken for such ulterior motives as 'edification' or the salvation of his soul. Since,

[1] Compare with the "bird's nest" effects produced by the child such a Surrealist picture as S. W. Hayter's 'Work in Progress' (Herbert Read, *Surrealism*, Plate 38).

[2] Compare the arrangement of objects in the 'Poème-objet' by the Surrealist, André Breton, illustrated in Herbert Read's *Surrealism*, Plate 13.

therefore, artistic activity consists in making things for fun and aesthetic pleasure consists in the fun of contemplating things made for fun, it follows that art has no place in the *Weltanschauung* of the puritan; and it is on this account that, if puritanism has one eloquent Milton to its credit, most of its Miltons are admittedly mute and inglorious.

V. Art and the dream

Having proceeded sufficiently far in our enquiry to reach some conclusion regarding the general nature of artistic activity, we must now examine in greater detail what artistic activity involves. At this point it will serve our purpose best to focus our attention upon the work of art and to attempt to reach some understanding of the psychological significance of the work of art in terms of the psychology of the artist. Our best prospect of reaching such an understanding seems to lie in comparing the work of art with another product of mental activity, about which we already know something, and to which the work of art is, without any question, closely related—viz. the dream. We know that Coleridge's 'Kubla Khan' was the embodiment of a dream in poetry; and, if a dream is not a work of art in itself, the recorded dream bears all the marks of a work of art, whether it be embodied in a drawing or simply in a written narrative.

As Freud has pointed out, the dream is the product of unconscious trends, which are repressed and thus denied not only direct expression in action, but also direct access to consciousness. The cognitive equivalents of the repressed urges consist in phantasies, which are likewise repressed. With the relaxation of repression which occurs in sleep there is a relative increase in the pressure of the repressed urges, which then mobilize sensory impressions and memory traces in such a way as to enable these to become the disguised expression of repressed phantasies. The resulting product is, of course, described by Freud as 'the manifest content' of the dream; and the mental processes, whereby the disguise is effected, are described as 'the dream work'. The work of art is distinguished from the dream above all by the fact that it is a product of motor activity. In the state of sleep motor activity is ordinarily precluded; and, consequently, it is impossible for the product of dream-work to be embodied in outer reality. If, however, it cannot be embodied in outer reality through the agency of motor activity, the product of dream-work is none the less exteriorized through the only alternative channel open to it. It is exteriorized as a sensory phenomenon. The manifest content of the dream thus acquires hallucinatory vividness and provides the dreamer with the

temporary illusion of experiencing outer reality. The work of art, on the other hand, is a product of motor activity and is thus enabled to find an actual place in outer reality, in contrast to the illusory reality of the dream-world. Apart from its embodiment in outer reality, however, the work of art bears a remarkably close analogy to the dream; and it is for this reason that we are conscious of no incongruity, when we see a mediaeval Gothic cathedral described in the guide-books as 'a dream in stone'.

According to Freud's theory of dreams, the primary function of the dream is to provide some means of expression for the repressed urges of the dreamer and, by relieving their pressure, to reduce psychical tension. We must regard the work of art as fulfilling a similar function in the mind of the artist. Freud has pointed out, of course, that, if the repressed urges received undisguised expression in the dream, sleep would be interrupted owing to the anxiety which would result from such a contravention of the standards accepted by the ego. Hence the necessity for dream-work, which enables a compromise to be reached between the satisfaction of the repressed urges and the need for sleep. Pursuing the analogy between the work of art and the dream, we must postulate a process of '*art-work*', which performs a function in relation to the work of art similar to that performed by dream-work in relation to the dream. Like dream-work, art-work must be regarded as essentially unconscious. It precedes the emergence of the creative phantasy into the artist's consciousness; and the creative phantasy which emerges is thus analogous to the 'manifest content' of the dream. It is through the agency of art-work that the repressed phantasies of the artist, in the form of 'manifest content', are placed at his conscious disposal for embodiment in works of art—and here it should be noted that the artist's conscious disposal of such manifest content is in turn analogous to the process of 'secondary elaboration' which Freud ascribes to the conscious mental activity of the dreamer. Art-work thus provides the means of reducing psychical tension in the artist's mind by enabling his repressed urges to obtain some outlet and satisfaction without unduly disturbing his equanimity; and it is in virtue of this function of art-work that we are justified in regarding the work of art as something that is made 'for fun'. It must be recognized, however that, just as dream-work can fail, so can art-work; and, when art-work fails, we may expect the equanimity of the artist to suffer. We may expect the quality of the resulting work of art to suffer also; for, since the character of the work of art is determined by art-work, the quality of the art-work must necessarily affect the quality of the work of art. That we are justified in our expectations may be illustrated by reference to the

work of the painter Sims, who committed suicide in a state of unsound mind in 1928. This artist's pictures showed progressively less evidence of art-work as his mental condition deteriorated; and the quality of his art suffered accordingly.

A further analogy between the work of art and the dream may be observed. In the case of the dream, the complexity of the dream-work depends upon the relative strength of the repressed urges and the factors responsible for repression. In the case of the work of art, the complexity of the art-work is similarly determined. This at once becomes obvious, when we compare, say, the Madonna by Leonardo da Vinci in the Hermitage at Leningrad with a picture entitled 'Maternity' by the Surrealist, Joan Miró.[1] In both pictures maternity is the theme; but, whereas the former shows considerable evidence of art-work, the latter, like most Surrealist works of art, shows comparatively little. It requires no very profound study of Surrealist works of art to convince us that the comparative poverty of the art-work which they display is directly related to pressure of unconscious phantasy combined with weakness of repression. For it is the avowed purpose of the Surrealist school to break down the barriers existing between the world of the unconscious and the world of outer reality; and it was in conformity with this purpose that in the recent Surrealist exhibition in London (1936) a place was found for examples of asylum art. Although it is hard to believe that all the criticism levelled against the Surrealists is based upon strictly aesthetic considerations, the comparative poverty of their art-work seems to justify the view that Surrealism does not provide us with art of a very high order. Our previous conclusion that anything that is made for fun is a work of art does not permit us to say that in the complete absence of repression there would be no art at all; but there can be no doubt that art-work, like dream-work, is dependent upon repression, and that without repression no high achievement in art is possible.

VI. Inspiration

Mention of the Surrealist school makes it now relevant to consider the nature of the unconscious urges in which the 'inspiration' of the artist has its source; and, in view of the unconscious nature of these urges, we must turn for our knowledge of them to such information as the psycho-analytical study of the unconscious provides.

Here let us remind ourselves that, according to psycho-analytic a

[1] This picture by Miró is illustrated in Herbert Read's *Surrealism*, Plate 54.

findings, human nature is governed by two great groups of incompatible urges springing from unconscious levels. The first group, described as 'the libido', consists of impulses embodying the life principle and expressing themselves characteristically in the fields of self-preservation, sex and love, whilst the second group consists of aggressive and destructive impulses, which represent a denial of the life principle. The coexistence in the individual of these two incompatible groups of impulses inevitably leads to a mental conflict of the most distressing kind; for in early childhood both the libidinal and the destructive impulses are in large measure directed simultaneously towards the same persons—either separately (i.e. ambivalently) or in a state of fusion (i.e. in the form of sadism). At the level of object-love, of course, the destruction of love-objects occasions intolerable guilt; but, even at the preceding narcissistic level, it occasions intolerable anxiety, since it involves the removal by the individual's own agency of his own sources of libidinal satisfaction. It is to deal with the anxiety and guilt engendered by destructive phantasies regarding love-objects that repression is originally instituted in childhood; and it is owing to the persistence of such phantasies in the unconscious that it is maintained in adult life. Nevertheless, whilst the rationale of repression lies in a need to safeguard the ego against the consequences of destructive urges, the intimate association of libidinal with destructive urges in early life inevitably results in a considerable repression of the libido.

Since artistic activity is essentially creative, we should expect art to be determined by exclusively libidinal urges; and it was in this light that art was originally regarded by psycho-analysts, before the importance of the destructive urges had been appreciated, and the significance of ambivalence and sadism had been properly understood. In the light of more recent psycho-analytical discoveries, however, it becomes evident that the destructive impulses must play an important part in the phenomenon of art. It becomes evident, for example, that art may be a channel for the expression of sadistic phantasies; and that this actually happens may be seen in the pictures of Goya and of the Surrealists.[1] Of course, the sadism of Goya and the Surrealists is expressed chiefly in the subject-matter of their pictures; but sadism may be expressed in the brushwork of a picture, even when it is absent in the subject—as in the case of Van Gogh.[2]

The influence of the destructive urges within the field of artistic

[1] *Vide* 'Tampaca' ('Desastres de la Guerra', No. 36) by Goya and 'Spanish Night' by the Surrealist, Francis Picabia (Herbert Read, *Surrealism*, Plate 76).

[2] Consider, for example, Van Gogh's 'The Cypress Tree'.

activity is by no means confined, however, to the embodiment of sadistic phantasies in works of art and the expression of sadistic impulses in their execution. No investigator has done more to enlighten us regarding the prevalence and strength of destructive phantasies in the unconscious than Melanie Klein; but Melanie Klein has also shown that these destructive phantasies are characteristically accompanied by compensatory phantasies of restitution. These phantasies of restitution arise as a means of alleviating the guilt and anxiety engendered by destructive phantasies. Their function is to provide some reassurance regarding the integrity of the threatened love-object; and, since the preservation and enhancement of its objects of attachment is the great concern of the libido, we must regard phantasies of restitution as libidinal manifestations in spite of the fact that they owe their origin to the presence of destructive urges. The cultural importance of phantasies of restitution can hardly be exaggerated; for, once the operation of the principle of restitution[1] has been recognized, it can be seen to underlie all the strivings of man after the perfect, the ideal and the absolute. It is this principle that spurs him on in his search after truth, goodness and beauty; and it is this principle that turns his steps towards religion. From the point of view of our present enquiry, of course, what concerns us particularly is the part played by phantasies of restitution in art. If art provides a channel of expression for sadistic phantasies, we have equal reason to believe that it provides a channel of expression for phantasies of restitution. Indeed, since the chief source of inner tension is found to lie in the pressure of destructive urges, and since artistic activity both relieves this inner tension and is essentially creative, we are justified in concluding that the principle of restitution is the governing principle in art.

Within a considerable area of the artistic field the dominance of the principle of restitution is easily appreciated. Greek art perhaps provides the most obvious example; for it is hard to imagine any more convincing attempt to establish the integrity of the object than that represented by the symmetry of Greek architecture and by the perfection of form and purity of line, which are such obvious features of Greek sculpture.[2] The work of the Dutch painter Vermeer also provides an obvious illustration

[1] The term 'principle' is not used in any mystical sense either here or in what follows. Psychologists who are suspicious of the phrase '*principle* of restitution' may, if they so desire, substitute '*motive* of restitution' without altering the writer's meaning. Cf. Freud's 'pleasure *principle*'.

[2] Consider, for example, the Apollo Sauroctonus (after Praxiteles) in the Louvre. Even the fragment of the Aphrodite (Zouaglia) in the Louvre conveys the impression of 'the integrity of the object'.

of the principle of restitution; for the tranquillity of his interiors and the meticulousness of his painting give us a very definite impression of the integrity of the object. These examples, of course, are drawn from art imbued with the so-called 'classical' spirit, in which the principle of restitution might be expected to be more obvious than in 'romantic' art. To convince ourselves, however, that the principle of restitution also reigns within the field of romantic art, we have only to reflect for a moment upon the atmosphere which pervades the great religious art of the Middle Ages in Europe.[1]

When we turn to some of the moderns, we find a very different state of things. In many pictures by such artists as Paul Klee and Picasso (even in his pre-Surrealist period) the objects represented give the impression of having been either grossly distorted or broken up into fragments—or else subjected to both these mutilating processes.[2] Pictures of this nature very naturally produce in the man in the street an impression of chaos; and they must be regarded as giving expression in no uncertain terms to that sadistic, 'tearing in pieces' tendency, to which Melanie Klein has drawn attention. Even in such pictures, however, we can find evidence of restitution—as when, for example, the parts of an apparently dismembered figure are found to have been reassembled, albeit in an unfamiliar way. Further, while such pictures may represent objects reduced to fragments, the ordered arrangement of these fragments in the composition must be regarded as a form of restitution.

The same combination of sadism and restitution may be illustrated in the case of the Surrealist painter Dali, in whose work the sadistic, 'tearing in pieces' tendency is often expressed with considerable licence. Take, for example, his picture 'Spectre du sex-appeal'.[3] Here we see the minute figure of a boy contemplating a colossal female figure, whose head merges into a rugged mountain mass in the background. The figure is deformed, contorted and mutilated; and various parts of the body are missing. Yet, apart from the unifying effect of the composition, evidences of restitution are not wholly lacking from the subject itself; for the figure is propped up by crutches and the missing parts of the trunk are at any rate replaced by sacks. In certain other pictures of Dali's the human body is represented as partly constituted by a tier of half-opened

[1] Consider, for example, the statues of 'Church' and 'Synagogue' to be seen at Strasbourg Cathedral.

[2] Consider, for example, Paul Klee's 'Dynamics of a Head' (Read, *Surrealism*, Plate 44) and Picasso's 'The Woman with the Golden Breasts' (Read, *Surrealism*, Plate 78).

[3] *Vide* Salvador Dali, 'La Conquête de l'Irrationel' (Editions Surréalistes, Paris), Plate 1.

drawers, out of some of which the contents are seen to dangle (*vide* Herbert Read, *Surrealism*, Plate 25). Here again we have an expression of the 'tearing in pieces' tendency. Yet the theme of restitution is not wholly absent; for, if a drawer may be opened for the removal of its contents, it may also be shut again with the contents replaced.

VII. Symbolism

These pictures in which part of the human body is represented as a chest of drawers bring us back to the question of art-work; and they direct our attention in particular to the part played by symbolism. The substitution of the conception of opening a drawer for that of tearing the contents out of a body obviously represents a certain amount of art-work. In spite of the sadism of the phantasy represented, the occurrence of such a substitution is evidence that repression is not altogether inoperative. It indicates that some degree of disguise is necessary to enable the artist to express his sadistic phantasies without undue guilt and anxiety. The art of Dali provides particularly favourable material for the study of art-work, because, when his pictures are surveyed as a whole, all the various stages of symbol-formation are found to be represented. If these various stages are placed in series, we obtain a bird's eye view of the process of symbol-formation. Thus in one of Dali's pictures we may find a disembowelled body, in another a body containing a tier of half-opened drawers and in yet another a chest of drawers detached from the body altogether. Having studied such a series, we are then in a position to state categorically that, when a chest of drawers appears in one of Dali's pictures, it represents a body.

Sometimes the emergence of the symbol may be studied within the scope of a single picture of Dali's, e.g. in the picture entitled 'Le sevrage du meuble aliment' (a title which may perhaps be translated as 'Being weaned from Nutrient Furniture').[1] The background of this picture is a land-locked bay; and on the shore in the foreground we see receding to the left a row of four boats with the bows pointing out to sea. In the fore-front of the picture, on the right of the boats and in series with them, is the seated figure of a woman, whose back is half-turned towards the spectator, and whose outstretched legs lie parallel to the boats. Her bowed head is swathed in a bandage; and a large rectangular block is missing from the trunk of her body, which is supported by a crutch. Just beyond the feet of the woman is a rectangular bedside-pedestal consisting of a cupboard surmounted by a drawer, which is half-open. A study of this

[1] *Vide* David Gascoigne, *A Short Survey of Surrealism*, p. 32.

picture makes it obvious that the rectangular pedestal represents the missing portion of the woman's body, which is also rectangular. Further, since the body of the woman is in series with the boats, it becomes obvious that the boats are equivalent to the bodies of women; and this is confirmed by the fact that the boat nearest to the woman has attached to it a cable corresponding to the crutch that supports the woman's body. From a study of this one picture we are thus enabled to infer that, when boats and bedside-pedestals occur in other pictures by Dali (as they frequently do), they represent a woman's body.

The interest of Dali for our immediate purpose lies in the fact that his art lends itself in a special degree to a study of the genesis of symbols under the influence of art-work. His pictures, like those of most Surrealists, cannot be regarded as representing a very high level of artistic achievement. There is, so to speak, a low coefficient of repression in relation to the strength of the unconscious urges expressed; and the result is that the art-work is comparatively meagre. It is for this very reason, however, that Dali's art is of such psychological interest. In the case of the more classical, as well as the more conventional, artists, the art-work has already reached such a stage of elaboration in the unconscious that, by the time the symbol emerges into the artist's consciousness, both its origin and its significance have become obscured. In the case of Botticelli's 'Birth of Venus' the title of the picture gives us a clue to the fact that the shell, above which Venus stands poised, must represent a maternal body. Otherwise the significance of the shell would be impossible to interpret in the absence of such special knowledge as the psychoanalytical study of dream-symbolism provides. It would be impossible to interpret, because Botticelli's art-work is so much more elaborate than that of Surrealist artists. It is only when art-work is relatively undeveloped that we are in a position to trace the origin of art-symbolism. This is just what we should expect on the analogy of dream-symbolism; for the knowledge of dream-symbolism which we owe to psycho-analysis is dependent on the fact that, as repression is overcome in the course of psycho-analytical treatment, dream-work becomes progressively simplified until a point is reached at which the significance of the symbol becomes apparent.

VIII. Art-work and the principle of restitution

It is now time for us to consider the relationship between art-work and the principle of restitution. Art-work, as we have seen, is analogous to dream-work; and, according to Freud's original formulation, dream-work

is the means whereby repressed phantasies are modified, or disguised, in such a way as to enable them to escape repression and to express themselves under hallucinatory form in a dream. Whilst this formulation must be accepted as correct so far as it goes, it can hardly be regarded as adequate, in that it leaves us in some doubt as to the source of the dream-work. It must be assumed, however, that Freud laid the *onus* of the dream-work upon the impulses investing the repressed phantasies, since, for explanatory purposes, he represented the operation of repression under the figure of a hypothetical 'censorship', which the repressed impulses could only elude by adopting a disguise. The conception of the 'censorship' was, of course, only a stop-gap conception pending further investigation of the agency of repression; and, when Freud subsequently formulated his theory of psychical structure, the function of the censorship was attributed to the 'ego-ideal' or 'super-ego'. In the light of Freud's eventual theory of psychical structure, the source of dream-work at once becomes clear. Dream-work is seen to be a function of the ego, which thereby modifies phantasies engendered by the instinctive id-impulses in deference to the demands of the super-ego. Once this conception is reached, it necessitates a significant change in Freud's original dream-theory; for it implies that dream-work is not merely a negative, but also a positive process. Dream-work ceases to be merely a means of disguise, whereby repressed impulses evade an internal censorship, and becomes also a positive gesture on the part of the ego towards the super-ego. Art-work must be regarded in a similar light. It is an activity of the ego performing a dual function in the economy of the artist's mind. On the one hand, art-work modifies repressed phantasies in such a way as to enable them to elude the vigilance of the super-ego and so to become available for embodiment in works of art—by this means affording the repressed impulses an opportunity for expression and relieving the *tension between the repressed impulses and the ego*. On the other hand, art-work enables the ego to convert phantasies unacceptable to the super-ego into positive tributes to its authority and so relieves the *tension existing between the ego and the super-ego*. Through its dual function, art-work is thus the means of producing that general relief of tension which the artist experiences, when he produces a work of art.

So far as the psycho-analytical study of art is concerned, attention has been hitherto largely concentrated on the first function of art-work; and this fact has given rise to the criticism that, while psycho-analysis may shed some light on the nature of the unconscious urges prompting artistic activity, the central problems of the psychology of art fall outside its

purview. This criticism ceases to apply, however, once the second function of art-work is envisaged. In virtue of this function, art is seen to be not only a sublimated expression of repressed urges, but also a means whereby *positive values* are created in the service of an ideal. The ideal served is the super-ego; and the creation of positive values is an act of restitution on the part of the ego.

We have seen, of course, that, even at the narcissistic level, phantasies of restitution occur as a means of relieving the anxiety engendered by destructive phantasies regarding love-objects. We must, therefore, envisage a form of art consisting simply in the embodiment of such restitutive phantasies; and the simpler and more primitive forms of art lend themselves to interpretation in this sense. We must so regard the earliest artistic efforts of the child and such examples of prehistoric art as consist simply in the outline of an animal or of a figure scratched upon a rock. Nevertheless, it is doubtful whether, either in the history of the individual or that of the race, art can be regarded as attaining a very high standard until the level of super-ego formation is reached. Once this level is reached, it is natural that restitution should take the form of a tribute to the super-ego; for the super-ego originates through an imaginative internalization or introjection of the love-objects threatened in destructive phantasies. Within the sphere of morals, the tension created by the pressure of the super-ego upon the ego is experienced as 'the voice of conscience'; and it is accompanied by an urge in the individual who experiences it to *do* something by way of restitution for destructive acts committed or destructive impulses harboured. Within the sphere of art, the tension created by the pressure of the super-ego is experienced as an itch to *make* something; and the creation of a work of art thus becomes a means of restitution analogous to the moral act.

At first sight it may not appear easy to relate the conclusions which have now been reached regarding the restitutive function of art to our previous conclusion that a work of art is something that is made for fun. It seems desirable, therefore, that some attempt should be made to clarify their relationship. As already indicated, the more primitive forms of art would appear to belong essentially to a level of development antecedent to that of super-ego formation. Since, therefore, at this level the work of art cannot be regarded as a tribute of restitution paid to the super-ego, it must be regarded simply as an embodiment of restitutive phantasies; and the satisfaction afforded by its creation must be attributed to a relief of the inner tension occasioned by repressed destructive impulses. In this connexion it is not without significance that, when the

young child is asked what his earliest and apparently incoherent drawings represent, it is usually found that they are intended to represent such objects of his destructive phantasies as 'Mummy', 'Daddy' and 'Nanny'. This fact accords with the restitutive nature of the child's artistic activity; but the further fact that the child's explanation is commonly accompanied by a laugh indicates that the drawings are at the same time made 'for fun'—the 'fun' being due, of course, to the relief of inner tension effected by the act of restitution. At this primitive level restitution would appear to be made to the object represented by the work of art itself; and in this connexion it is interesting to note not only that such art is characteristically representational in spirit, but also that the earliest examples of prehistoric art consist for the most part in representations of animals hunted by primitive man (cf. the theory that such representations had the magical intention of promoting the fertility of the animals in question). Once the level of super-ego formation has been reached, it is only natural that the restitution effected by artistic activity should be made to the super-ego instead of to external objects; for the replacement of external by internal objects has been shown by Freud to constitute the essence of super-ego formation. This substitution does not, however, affect the essentially restitutive nature of artistic activity. Neither does it affect the essential nature of the satisfaction which artistic activity provides. For, in so far as the act of restitution is determined by a need for the relief of inner tension, the creation of a work of art must be regarded as 'making something for fun'.

Once we come to recognize the essentially restitutive nature of artistic activity, we find ourselves in possession of a clue to a number of further problems, among which the vexed question of the artist's audience is by no means the least important. At an earlier stage in the present paper an attempt was made to show that the essence of art lies in the creative act of the artist quite irrespective of any effect produced by the work of art in the percipient, and that, in so far as this is so, the artist requires no audience. This conclusion must now be qualified by the admission that artistic creation does seem to require an audience *of a kind* after all. The audience for which the artist creates is not, however, the conventional percipient. It consists in the objects to whom restitution is made, whether, as in the case of primitive art, these be external objects (actual or phantasied), or, as in the case of more advanced art, they be the internal objects represented in the artist's own super-ego.

ANOTHER APPROACH TO THE PROBLEM OF ACCIDENT CAUSATION

By H. BANISTER

From the Cambridge Psychological Laboratory

I. *Reasons for the enquiry* (pp. 304–305).
II. *Material* (p. 305).
III. *Method of enquiry* (p. 306).
IV. *Results* (pp. 306–307).
V. *Discussion* (pp. 307–313).
(1) *Stupidity* (pp. 307–308).
(2) *Foolhardiness* (pp. 308–309).
(3) *Carelessness* (pp. 309–310).
(4) *Distraction* (p. 310).
(5) *Hurry* (pp. 310–311).
(6) *Anger and annoyance* (p. 311).
(7) *Worry* (p. 311).
(8) *Wish fulfilment* (pp. 311–312).
(9) *The first accident* (pp. 312–313).
VI. *Summary* (pp. 313–314).

I. Reasons for the enquiry

In spite of all the efforts which have been made in recent years to reduce and to control the numbers of accidents, these still entail an enormous loss to the community, particularly on the road and in industry.

The earliest attempts to cope with this problem were by legislation, and in this way much has been accomplished. It seems doubtful whether much further alleviation can be brought about by this means. Now, for some years, the industrial psychologist has been attacking the problem with success by introducing various modifications into the working conditions and by vocational selection; while further progress has been made through the appreciation of the fact that some individuals are more liable to accidents than are others, and that these individuals can, to a certain extent, be picked out by specific tests, principally of motor co-ordination

temperament and intelligence.[1] But these are statistical results and, like most statistical results, are not valid when applied to any specific individual, because, presumably, other factors, not discovered by the tests affect susceptibility to accidents.

It seemed worth while, therefore, to enquire into the causes of a number of accidents from another point of view, taking each accident separately and endeavouring to find out the cause in a friendly personal interview.

II. Material

The enquiry was carried out at the Papworth Village Settlement. Papworth consists of hospitals for adult tuberculous patients with hostels and houses for those who are convalescent or in whom the disease has been arrested. In the houses ex-patients live normal lives with their wives and families.

Occupation is found for the ex-patients and for the adult members of their families, as well as for those who are convalescent, in the various workshops, where most kinds of trades are carried on, from mending boots and shoes, carpentry and building to upholstery, trunk making, bookbinding and printing. In spite of the fact that the work-people are, with a few exceptions, physically subnormal so that they could not compete in the ordinary industrial market, the work is carried on under ordinary healthy industrial conditions. The one great difference from work in the outside world is that the work-people are under constant and careful medical supervision, so that the hours and kinds of work are graded to each individual's strength, while at the same time there is no fear of unemployment if a worker stays away for a day or two. The result is that the attitude of the workers to their work is healthier than would normally be found outside. The work is not, as it so often may be in industry, something irksome which the workers must do in order to earn a living; it is something they do because then they *are* living. Every patient looks forward to the time when he or she can work. It is the work which gives them hope and which is one of the principal factors in their treatment. In consequence many of the problems of industrial accidents do not arise here. There is no increase in the incidence of accidents towards the ends of the work shifts, there is no increase in incidence on Mondays as compared with other days, or after holidays, etc. But with such exceptions the psychological causes of accidents remain much the same at Papworth as elsewhere.

[1] Cf. Farmer, E. (1932), *The Causes of Accidents.* London: Pitman.

III. Method of enquiry

Workers who have accidents report at the surgery where their wounds are dressed. These workers were, sometimes on the same day, sometimes after an interval of a day or two, interviewed by the writer who, as a member of the honorary staff of the settlement, often already knew them intimately. In consequence the patient was, as a rule, willing to discuss the cause of his accident freely without feeling any need to suppress relevant information. At the interview, besides discussing the accident, the worker's family history was gone into; his attitudes to his work, to other workers and to life in general were considered; his day-dreams, night-dreams and worries were discussed, etc. In this way an idea of the psychological make-up of each individual was obtained.

This knowledge was further amplified by reference to each worker's medical history and by information given by the medical staff who knew each individual intimately in a way which is only possible in an institution of this kind. The whole of this material was then considered in relation to the specific accident.

IV. Results

Fifty-four accidents were investigated in this manner. Of this number forty-five occurred at work and nine during non-working hours.

A preliminary classification shows that the accidents were due to the following causes:

Environmental conditions	17
Carelessness on some one else's part	5
Inexperience	3
Stupidity	3
Foolhardiness	5
Carelessness of injured person and other psychological causes	21
Total	54

It will be seen that 32 % of the accidents are ascribed to environmental conditions and 68 % to other causes.

Those accidents which are due to environmental conditions, to carelessness on the part of some person other than the one injured, and to inexperience, will not be considered further here. Though they are of interest, in that an analysis of them indicates ways in which similar accidents may be avoided in the future, analysis of them throws little light on the reasons why expert workers have accidents for which they

are themselves responsible. This information is to be obtained by a consideration of the other accidents, viz. those ascribed in the first place to stupidity, foolhardiness, carelessness on the part of the injured person and to other psychological causes.

V. Discussion

The causes of the accidents which are to be discussed have been ascribed to stupidity, carelessness, foolhardiness and to other, so far unspecified, psychological causes. It is necessary that these various causes should be considered in further detail.

(1) *Stupidity.* Stupidity, as used here, does not readily denote any specific psychological quality. It does not denote that the individual would make a low score on an intelligence test, or that he would be considered intellectually below the average of those who do similar work efficiently. Again it does not signify that he is not paying attention to his work; if the accident were due to lack of attention it would be more properly placed in the carelessness class. Also it does not mean that he is in a hurry or is annoyed or that there is any other such reason why the accident happened. When an accident is ascribed to stupidity, it means that the worker was attending to his work, and that he was fully competent to do it, but that he has done something which he would not have done had he stayed a moment to consider some particular point in what he was doing, and that the reason why he did not stay the necessary moment to think was not that he was pressed for time or upset or anything else of that nature. He simply did not stop and think. He probably did not realize there was anything to be considered. There was some novel feature in the situation, the significance of which escaped him, or which he left to chance or attacked in an automatic manner, not realizing the need for consideration. Let me give an example:

Two members of a 'gang' of six were absent. The gang usually worked together when cumbersome goods had to be loaded or unloaded, and were skilled at the work which is intricate and demands nice judgement and delicate adjustments, with mutual understanding between the individual members of the gang. That this understanding, or bond of primitive comradeship, as Bartlett calls it, was present is evidenced by the fact that the men spoke of themselves as "*the* gang". On this particular day when two of the gang were away or employed elsewhere, concrete slabs had to be loaded on to a lorry, and, when loading such slabs, it is usual for three men to go to each end. The obvious allocation of the men, had anyone stayed to think, would have placed two skilled men with an unskilled

man at either end. But the thinking did not take place, as it undoubtedly would if the two had been new members of, say, a football side. The allocation was left to chance, and other psychological factors came in which determined, probably quite unwittingly, the arrangement of the gang. Although all the men knew each other, they formed in this particular situation two separate and distinct groups—the skilled, the old members of '*the* gang', and the unskilled—with the result which could have been foretold, one which is found over and over again in all such social situations: the new, unskilled men stuck together. They went to one end with one of the skilled men. But they did not know intuitively what to do; they were unable to appreciate the hints given by the others. As the injured man said: "They did not understand our methods and pushed at the wrong time." The result was a very badly bruised chest which kept the one who was hurt away from his work for a fortnight.

(2) *Foolhardiness.* Foolhardiness differs widely from stupidity, as defined above, though foolhardiness may in popular speech be described as stupid.

Psychologically foolhardiness is comparatively simple. The individual is attending to what he is doing, he realizes there is the risk of an accident but, sometimes out of sheer bravado, sometimes out of mere laziness, and perhaps more frequently in industry from a combination of these two reasons, he does what he well knows is taking a foolish risk, and there is an accident.

Many of the greatest joys in life are associated with the thrill of taking risks, usually quite needlessly. The child, the youth and the young man alike crave for the thrill of danger successfully encountered—the only games and sports 'worth while' to many are those in which there is danger. At the same time there is the, usually but half realized and wholly illogical, belief that though accidents do happen in such circumstances 'they will not happen to me'. As the individual grows older he appreciates the hard facts more clearly. He takes risks in a more organized manner where the risk run is realized, and pits his skill against a partially hostile environment over which he believes he can triumph. In his decline he is unwilling to take any risk and the greater part of the *joie de vivre* has departed. It is therefore to be expected that accidents due to foolhardiness are to be found chiefly among the young. It is an interesting and probably significant fact that, of the four workers under eighteen who attended the surgery on account of injuries sustained in accidents during the time of the enquiry, three were involved in accidents which were due to foolhardiness. The following is a typical example:

A man wished to remove some bits of wood from the side of a moving band-saw. He should have used a piece of wood for this purpose, but instead he tried to remove them with his fingers and, to make matters worse, instead of walking round to the far side of the saw when the risk would have been minimized, he stretched across the sharp edge of the saw, got too close and was lucky not to have lost two of his fingers instead of merely having them badly cut.

(3) *Carelessness.* Carelessness is an omnibus term which indicates that the worker has not been exercising due care in the performance of his task, and 'due care' is usually taken as a synonym for 'properly directed attention', with the further implication that 'properly directed' means 'concentrated'. In this form the proposition, that when accidents are ascribed to carelessness they are due to lack of concentration, is not strictly true. Hence it is worth while considering what is the real significance of 'properly directed attention'.

Since the work is being done in an environment, one cannot decide what is properly directed attention without considering both the work which is being done and the environment other than the work. Taking the work first, this may vary in industry from a repetitive task of a purely automatic kind to work requiring nice judgement with practically no repetition. For maximum efficiency the kind of attention required in these two types of work is different. For purely routine automatic actions the maximum of efficiency results when the individual goes through the movements in a purely routine automatic manner, when he becomes, as nearly as is possible, an automaton, with few conscious thoughts about his work. The ideal state is perhaps one of *reverie* in which he is attending to pleasantly toned day-dreams. He can then do his work happily while his thoughts are far away. The other type of work, with variety in its execution, requires on the other hand close concentrated attention all the time.

Again, when we consider the environment, it is obvious that different types of attention to this environment are necessary according as it is safe or otherwise. If it is safe the individual may move about at his work, as may be necessary, with impunity without attending to the environment at all; but, if the environment is not safe, it also must receive attention. In this latter case a more distributed type of attention will be an advantage. So it follows that properly directed attention may be of many kinds—attention which is properly directed in one environment or with one type of work is improperly directed in another type of environment or with another kind of work.

A number of accidents were due to improperly directed attention considering the type of work and the nature of the environment. One man, on a 'varied' job, was returning to his bench from the store with some material he required for his work. He was examining the material as he returned and in consequence did not notice the end of a cramp which was projecting from another bench. The environment was not safe enough for such concentration. Another man, when discussing his accident afterwards, said: "My mind wanders off sometimes. It has done this for about three years. It may happen when I am playing or when I am working. Then it goes off on to what I shall do when I am free, or to 'what might have been'. Sometimes—I remember one time when I was making a sliding door for a chicken hutch—it runs off and I 'come to' to find I have finished the job, yet I have no recollection of doing it. I am interested in languages and history, and have been studying Latin, Danish and Hindustani." This man had at one time been working on a machine for which he was obviously quite unsuited. After a short time, in consequence of accidents, he was put on to less dangerous work, but he still, as he put it, frequently "hit the wrong nail".

It is possible that, in the above cases, the differences between the individuals, which determine the type of work for which they are suitable and the environment in which they are safe, are temperamental, and so fairly constant.

(4) *Distraction.* Since, as has been pointed out above, properly directed attention is necessary if there are to be no accidents, it follows that any thing which changes the direction of attention, even if for only a very short time, is liable to cause accidents to workers who are in an environment and at work of a nature to which they are temperamentally unsuited. Dreamers and those who normally concentrate closely on their work will not, as a rule, be affected so easily as will the workers at neither of these extremes, by distractions, i.e. by changes in the environment which affect for a short time the attention the worker is giving to his work. The appearance of strangers, being spoken to by a neighbour, etc., if they disturb the attention at particular moments, are frequent causes of accidents. One man, a carpenter, was sharpening his pencil with a chisel "when a chap said something to me and I looked up". He cut a piece out of his finger.

(5) *Hurry.* In some circumstances the direction of attention may, for some specific reason, be changed almost completely and sometimes for comparatively long periods. Then an accident is to be expected even though the individual is working at a task and in an environment for

which he is normally peculiarly adapted. The chief causes of these comparatively long changes in the direction of attention appear to be hurry, anger and annoyance, and worry whether normal or pathological. These causes will be considered separately.

Hurry must be differentiated from speed. A task may be done quickly without there being any suggestion of hurry. Hurry implies that the individual is not getting along as quickly as he would; he is held up; he *must* get it done. The whole experience may be strongly and unpleasantly toned. The attention then becomes concentrated on the end, the means do not matter and, the more the need for hurry, the greater the risks taken so that, in extreme cases, as for example when one is rushing to the station to catch a train which is due to start, and the missing of which will involve much inconvenience and perhaps loss, one is, temporarily, properly to be classed with the foolhardy. It is not without significance that two of the four accidents due to hurry occurred during non-working hours.

(6) *Anger and annoyance.* There is no need to analyse the dangers of anger and annoyance. Not only is attention then distracted from the work to the cause of the anger or annoyance, but the nicety of adjustment is also upset. Excessive force is used, movements are made too violently and too rapidly, dangers are ignored and risks are taken, so that these individuals, too, tend to move into the foolhardy class until their anger disappears.

It is proper, of course, to maintain the distinction between the foolhardy, the hurried and the angry, but they are essentially the same; the one takes risks principally, perhaps, because of temperamental qualities, the hurried person because he is in a hurry, and the angry one because he does not care what happens. These workers would probably pass almost any test for muscular co-ordination, so that they would not fall into the accident-prone class of the industrial psychologist, yet they together account for over 25 % of the accidents considered in this paper.

(7) *Worry.* Worry produces its disturbing effects in a different way. The attention is distracted and the co-ordinations are upset, but the individual does not move into the foolhardy class. He is one who now would probably fail if tested for accident-proneness, and, at least for so long as the worry remains unresolved, is a source of danger to himself and, if circumstances permit, to others. It is of interest to note that there was only one such case among those analysed here, and this man was not a patient.

(8) *Wish fulfilment.* We come now to a quite different class, one in

which the accidents are brought about by more direct and probably unwitting if not unconscious, psychological processes, yet a class which would ordinarily be included under the general heading of carelessness. The difference between these accidents and those already considered lies in the fact that these accidents appear to be purposive while the others, except perhaps to the Freudian, are not. Two accidents fall into this category and they are sufficiently different in character to warrant a brief reference to each.

The first accident befell a discontented patient. At Papworth the workshops are closed for a week at the beginning of August when those who are well enough go away for a holiday. This particular man went away and did not want to return, not because he was dissatisfied with Papworth or with his treatment, but because he was dissatisfied with his work. It did not give him 'sufficient scope'. He wanted to do much bigger things. He could not stand the work—as a matter of fact he probably could not happily settle down to any work, for his discontent with the work was a projection of his own inner discontent. The result was that within half an hour of work beginning again he had an accident.[1]

In the second instance the man was helping to move an incubator. As he was bending down to take hold of it he remembered he had been told not to lift heavy weights, but he decided he 'ought to help' on this particular occasion. Such a conflict and attitude are typical of many psychologically determined symptoms in the psycho-neuroses and accidents in industry. There is no cause for surprise, therefore, that he fell and hurt his knuckle and did not lift the incubator.

(9) *The first accident.* There was only one instance during the three months of the enquiry of an individual having more than one accident of which he was himself the cause, and this man had three accidents between 3 o'clock one afternoon and 11 o'clock the following morning. They were all accidents which would ordinarily be ascribed to carelessness. For some reason, which is not at all obvious, his co-ordinations were upset. It may have been that the first accident, when he hit his finger with a hammer, was due to the distraction of his attention for some one of the reasons which have been considered above, and that then 'things just went wrong', as they so easily may. One often hears "My fingers are all thumbs to-day", and many workers in laboratories know how nothing will go right with their apparatus on certain days. Some small accident happens because, possibly, of some upset which has nothing to do with

[1] It is of interest to note that after the holiday no workman, with this one exception, attended the hospital on account of an accident for 9 days.

the work. Annoyance supervenes, force is wrongly adjusted and something else goes wrong, and so the accidents pile up during the day. At the same time this press of accidents does not occur to everyone, so it is perhaps dependent on some temperamental factor. But it is doubtful if this is what happened in this case, for there was a night's rest in between the accidents. It seems more probable that the man was maladjusted, if only temporarily; and this probability is borne out by his behaviour. He was fidgety and jerky in his movements, and uncommunicative in a way which is strong evidence for the presumption that something was hidden away.

VI. Summary

In this enquiry an attempt has been made to analyse out the psychological causes of the accidents which occurred at the Papworth Village Settlement during July, August and September 1936.

The chief causes were found to be:

(*a*) Stupidity. The worker meets some slightly novel situation in a maladaptive way because he does not pause to consider how it is best to proceed in this new situation. He is attending to his work which he is quite competent to perform.

(*b*) Foolhardiness. This is probably a temperamental characteristic and is more common among the younger workers. The foolhardy person attends to his work and realizes that it is dangerous to perform a certain action in a particular way, yet he takes the risk partly, perhaps, out of bravado and partly because he is too lazy to perform the action in the safe way.

(*c*) Improperly adjusted attention. There is no one particular type of attention which is suitable to all conditions of work. At one extreme purely automatic tasks require the withdrawal of attention so that the worker works oblivious of his task; at the other extreme many skilled tasks need very concentrated attention, while in between are the majority of tasks which require various amounts of concentration and withdrawal of attention. Again the environmental conditions must be taken into account. Unless the environment is uniformly safe both the extremes of attention may entail accidents.

(*d*) Distraction. Accidents are liable to occur when a worker's attention is taken away from his work through some change in his environment. These danger periods are usually of comparatively short duration.

(*e*) Hurry. When the whole attention is concentrated on speed the individual tends, temporarily, to go into the foolhardy class and accidents

are to be expected. Hurry is more liable to cause accidents during non-working hours than under industrial conditions.

(*f*) Anger and annoyance. With the emotional upset due to anger the delicacy of co-ordination tends to disappear, the direction of the attention is changed and the angry person is, for the time, to be included in the foolhardy class.

(*g*) Worry. Worry, besides distracting the attention, usually entails loss of sleep and a general upset which tend to disturb the co-ordinations and thus to make an otherwise efficient worker peculiarly liable to accidents.

(*h*) Wish fulfilment. In particular cases accidents occur when the individual wishes to avoid the work. This does not occur consciously.

(*i*) The first accident. One accident may be followed by others to the same person. This is probably due to some temperamental peculiarity.

It is to be noted that the ordinary tests for accident-proneness scarcely touch the fringe of these psychologically determined causes of accidents. It is possible for the very best performer at the tests for accident-proneness to be stupid (in the sense in which that word is used here), foolhardy, distracted, hurried, angry or annoyed, worried, wishful to avoid work, etc., and so to be, for a longer or shorter period, accident-prone. That is probably the reason why correlations between the test results and accidents are so low.

It would seem that this line of approach to the problem of accidents is worthy of being followed up.

In conclusion I would record my appreciation of the help I have received from Dr Stott and Dr Macullam.

(*Manuscript received* 10 *October* 1936)

THE WORK CURVE AND THE WORK DECREMENT

By J. A. BIRTWISTLE

I. *Introduction* (pp. 315–316).
II. *A new method of plotting the work curve* (p. 316).
III. *The line of 'best fit' to the work curve* (pp. 316–318).
IV. *Two new measures of work decrement* (pp. 318–320).
V. *Summary* (p. 320).
References (p. 320).

I. Introduction

The work curve derived from continuous mental work in general manifests a decrement in efficiency of output which traditionally has been measured by some such means as the difference between the first and final, or that between the maximum and minimum, outputs.

During the course of a research (1), tests in mental arithmetic which involved the addition, subtraction, multiplication and division of two figures were given to a group of London Senior School children at intervals over a period of nearly five months. The tests varied in length from 10 to 30 min., the size of the class ranged from 32 to 38 boys and girls of average age 13 years.

On comparison, the work decrements, derived from the work curves of the subjects as a group, were found to be conflicting and even contradictory, and this led to a desire for a more accurate measure of work decrement than the ones already employed. The first point of a curve was always found to be very high, and, together with the occurrence of a single deep trough, resulted in measures of decrement which obviously were not characteristic of the whole of the work curve.

It was thought that some measure was needed which would express the gradual and continuous decrease in efficiency of the whole of the work. This would be more characteristic of the slope of the work curve than any measure given by only two points on the curve, for such a measure is susceptible to error as the points themselves are liable to be the result of chance factors.

It was considered that the thing required was, if possible, a straight line which would be representative of the work curve, so that its slope might provide a more valid measure of work decrement. The difficulty of arriving at a satisfactory line of 'best fit' to the work curve proved to be great owing to the large fluctuations in the curve. The problem then, was to find a way of reducing the curve of output to approximately linear form.

II. A new method of plotting the work curve

A technique for plotting the work curve, new as far as is known to the writer, was employed for the first time. It originated in a suggestion by Philpott (2) to the effect that, from empirical results obtained by him, the curve of output has a general form roughly approximating to a rectangular hyperbola, and that, if the logarithms of time and output are plotted, the work curve "should tend to appear as a straight line, the fatigue decrement thus being measurable in terms of a simple linear expression".

The rectangular hyperbola can be expressed by the equation

$$xy = k \text{ (a constant)},$$

and taking logarithms

$$\log x + \log y = k' \text{ (a constant)}.$$

Hence if $\log y$ be plotted against $\log x$ a straight line will result and it will have a negative gradient of unity.

If the curve is a perfect rectangular hyperbola, by plotting log output against log time it will be translated into the straight line form. Actually, of course, this is not so, but if the curve of output does fairly approximate in shape to such a hyperbola, the plotting of the logarithms of time and output should effect a considerable flattening.

Following up this suggestion, the curve of a given test was now obtained by plotting the output vertically and the time horizontally on two-way logarithm graph paper, which is equivalent to plotting the logarithms of the outputs and time intervals. The resulting curves were flattened and smoothed to a considerable degree. As an illustration a typical work curve is given in Fig. 1, first in its usual form and secondly when two-way logarithm graph paper is used.

III. The line of 'best fit' to the work curve

The next step was to find lines of 'best fit' to the curves, for as noted above, the slopes of such lines should provide more satisfactory measures of work decrement than the traditional measures.

The larger the number of subjects contributing to the output totals, the less erratic is the resulting curve. When there was a comparatively large number of contributions, the work curve when plotted in this new manner was almost linear, except for the first point on the curve, which, with perhaps only one or two exceptions in 49 curves, was always well above the line of 'best fit' to the other points on the curve. Consequently in obtaining a line of 'best fit' the first point on a curve was ignored for the purpose of finding a work decrement measure.

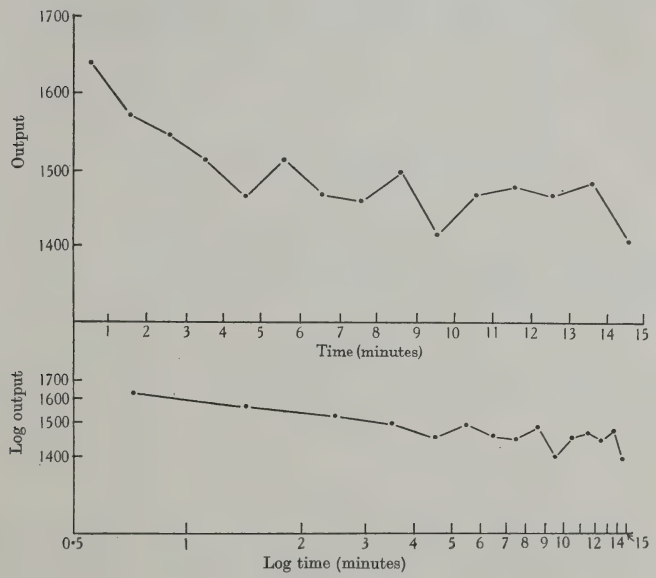


Fig. 1. Two forms of a work curve.

With curves considerably smoothed and flattened the best fitting line could be found comparatively easily, but with curves which showed more fluctuation a convenient and accurate method was by using one of the regression lines to the points on the curve.

Referring to Yule (3), p. 169), the line suitable for the purpose is

$$y = b_2 . x, \text{ where } b_2 = r \frac{\sigma_y}{\sigma_x},$$

and for a work curve x will be a time measure and y an output ordinate. Further, b_2 is the tangent of the angle of inclination of the line of 'best fit' and could serve as a satisfactory measure of decrement.

In the diagram shown in Fig. 2 of a hypothetical work curve of 10 min. duration, nine points are plotted. On a base line AB a convenient number of divisions are made each q mm. in length, and OO' is the mean height of the ordinates which meet the curve in the points l, m, n, etc. With axes

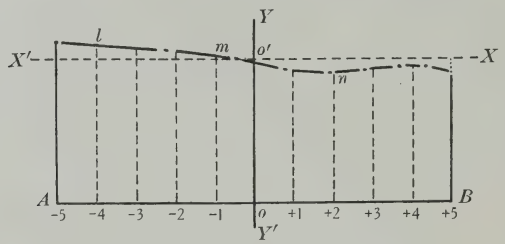


Fig. 2. Diagram of a hypothetical work curve.

XOX' and YOY' if the y values are measured in millimetres and the x values in the arbitrary units, the required equation for the regression line is

$$y = \frac{\Sigma xy}{q\Sigma x^2}.x,$$

and therefore the gradient of the best fitting line is

$$\frac{\Sigma xy}{q\Sigma x^2} = b_2.$$

IV. Two new measures of work decrement

The first measure of decrement made possible by plotting the work curve on two-way log paper (see Fig. 3 in which P is the first point on the work curve and RS is the line of 'best fit' to the curve) is the difference in output corresponding to the two end-points of the best fitting line. This difference in output (d), corresponding to the length of RQ, is readily found by reference to the vertical scale used on the graph paper.

This 'score' measure, as it may be called, is an absolute measure and is merely the refined equivalent of the usual measure obtained by taking the difference between the first and last minute's work. It is a refinement of the older measure because, whereas that deals with the actual outputs during these two periods of work, the 'score' measure is derived from outputs represented by the end-points of the line of 'best fit'.

Being obtained from the end-points of a straight line which is characteristic of the work curve, it is representative of the decrease in efficiency of the work during the period of testing, whereas the usual measure (being derived from a hyperbolic curve) is not. Moreover, unlike the latter, it is not liable to be influenced by chance factors which may affect the final output, nor is it magnified by the relatively great output which generally occurs in the first period of working.

The score measure, however, like the usual absolute measure, is dependent on the time interval used in plotting the work curve, for the fall in score from first to last point is greater when the interval is long.

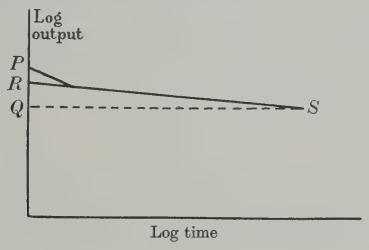


Fig. 3. The general form of the work curve.

A variation of the score measure becomes desirable when comparing the decrements of curves, the mean output of which differs considerably. It is similar to the percentage fatigue measures commonly employed in that it is a relative measure and is the ratio of the total decrement to the total score (or mean score m) in the combined work of the subjects. It is

$$\frac{d}{m} \times 100\ \%.$$

This second or 'percentage score' decrement is a refinement on the score measure because it is affected to a less degree by the output of work, and it is also an improvement on the ordinary percentage fatigue measure for the same reasons that the score measure is superior to the ordinary absolute measure of decrement. It is unsatisfactory in some respects just as is the usual relative measure of decrement, for its magnitude depends on

(i) the time interval used in plotting the work curve, for the interval will determine the time at which the first point occurs;

(ii) the duration of the work, which will determine the end-point of the best fitting line.

The third measure of decrement which overcomes these limitations is the gradient of the line of 'best fit', RS, viz. $\frac{RQ}{QS}$, in which RQ and QS are measured in units of length, and it is identical with $b_2 = r\frac{\sigma_y}{\sigma_x}$ of the regression equation.

This gradient measure of decrement is similar to the percentage score measure in that it is a relative measure. The difference RQ in the lengths of the ordinates of the best fitting line enters into the tangent of the angle of slope, and being the difference of two logarithms, it is the logarithm of a ratio.

In actual practice it was found that these two relative measures of decrement correlated very highly (0·96 for 49 curves). Either measure can be converted to the alternative form by reference to the best fitting line to the work curve when graphed as stated.

V. Summary

Summing up the whole procedure: While the usual methods of measuring work decrement may often hold, another method of plotting the curve of output may be expedient when difficulty arises. In this method the work outputs are plotted against the time intervals on two-way logarithm graph paper in order to reduce the curve to linear form. The line of 'best fit' is obtained by means of one of the regression equations, and from this, an absolute score measure and two relative measures of work decrement may be derived.

REFERENCES

(1) Birtwistle, J. A. (1936). "A comparative study of the results of homogeneous and heterogeneous work in mental arithmetic." Unpublished M.A. Thesis, Institute of Education, University of London.

(2) Philpott, S. J. F. (1932). "Fluctuations in human output." This *Journal*, xvii, 112.

(3) Yule, U. (1932). *An Introduction to the Theory of Statistics*. London: Griffin.

(*Manuscript received* 28 *March* 1937)

FOURTH REPORT ON A LAMARCKIAN EXPERIMENT

By WILLIAM McDOUGALL

Part I. *Continuation of previous experiments* (pp. 321–334).
(1) *Introduction* (pp. 321–322).
(2) *Résumé of previous work* (pp. 322–323).
(3) *Continuation of the experiment during the fourth period* (pp. 323–330).
(4) *The adversely and the favourably selected series of generations trained in the tank* (pp. 330–333).
(5) *Further cross-breeding experiments* (pp. 333–334).
Part II. *Reply to some criticisms of the First, Second and Third Reports* (pp. 335–345).
(1) *Dr Crew's criticism* (p. 335).
(2) *Crew's use of an easier task* (pp. 335–338).
(3) *The variability of errors in the control groups* (pp. 338–345).
References (p. 345).

Part I. Continuation of previous experiments

(1) *Introduction*

This article falls naturally into three parts. Part I reports the continuation of experiments described in three preceding articles[1](1, 2, 3). Parts II and III[2] are concerned with the examination of various criticisms made upon the content of those articles, and, especially, with certain counter-criticisms of the extensive experiment designed and conducted by Dr F. A. E. Crew as a basis for criticism of my work. It was largely because I was aware of this work proceeding in Dr Crew's laboratory at Edinburgh that I have so long postponed making any thorough reply to criticisms: for the first publication reporting in full Dr Crew's work was his article of August 1936(4). But also I have postponed replying at length to

[1] Since the name of Dr J. B. Rhine appeared on my third report as co-author and collaborator, I must explain here that he has taken no active part in the work of this fourth period of the experiment because his time and energy have been wholly absorbed in the direction and supervision of the work of the large sub-department of parapsychology which he has developed within my department at Duke University. Dr Rhine allows me to say that his belief in the value and promise of this experiment is in no way diminished.

[2] Part III will be published in a subsequent number of this *Journal*.

criticisms because I hoped that the best answer to many of them would be given by the progress of my experiment, especially those branches of it designed and undertaken with a view to meeting criticisms and interpretations offered as alternatives to the Lamarckian hypothesis; an expectation which, I think, the present report in large measure justifies.

(2) *Résumé of previous work*

It seems likely that the present report may be read by some who have not seen the three earlier reports. I crave pardon, therefore, of those who have read those three reports, if I repeat in condensed fashion much of their substance.

The experimental work of the period since June 1933, the date of the Third Report (a period called 'the fourth period' extending from June 1933 to April 1937), has consisted mainly in the continuation through further generations of the training of albino rats by my 'tank-procedure'. This procedure together with some variations of it was described in detail in my third report. Here I cannot do better than cite Dr Crew's very clear brief description of it. "McDougall trained rats by a standardized procedure to avoid the brightly lit and electrified route out of a water tank. The rat is lowered by hand gently into the middle blind compartment of the tank out of which there are two lateral ways leading to platforms, one of these, alternately on the right and left, being illuminated and so wired that the rat, stepping on to it out of the water, receives an electric shock. The tank is such that the rat has the choice of leaving the water either by way of the dim platform and thence out of the tank without receiving a shock, or else by way of the lit platform and getting a shock.

"In order to accustom the rat to the conditions of the experiment, each is given six runs in the tank, with the light alternating but without shock, as a preliminary to the actual training period during which each rat is placed in the tank six times daily.[1] On the first of these occasions the light and shock are on the right of the rat as it swims down the middle blind compartment, on the second they are on the left, on the next they are on the right, and so on; and this practice is continued until the rat has mastered the task; that is, until it has left the tank twelve times

[1] It is perhaps worth while to point out that this shockless series of six immersions on the first day (hereafter called 'zero-day') was adopted as part of the routine procedure, not only in order to accustom the rats to the conditions of the experiment, but also and chiefly because I expected that any Lamarckian transmission of the effects of training, if it should occur, would reveal itself as an increasing preference for the dim route, *D*, as against the bright route, *B*, in successive generations: an expectation which seems after many years to be in course of realization. Cp. Tables VII and VIII.

in succession by the dim unshocked route. The training is completed when the rat 'learns to discriminate between the bright and the dim gangway (in anthropomorphic terms it learns to accept the bright light as a signal warning it of a shock)'."

(3) *Continuation of the experiment during the fourth period*

The general course and effects of such training prolonged through many generations is shown in Table VII which reproduces and brings up to date (April 1937) Table II of the third report, and also combines with it the relevant figures of Table V.[1]

Table VII

Generation of T.R. stock	No. of rats	Average no. of errors per rat	No. of errors made by Best rat	No. of errors made by Worst rat	Zero-day choices *D*—*B*
13	23	68+	30 (1)—	90+	76—62
14	10	80	42 (1)—	102	31—29
15	10	70	39 (1)—	96	24—36
16	5	73	39 (1)—	88	14—16
17	11	46	9 (1)—	147	35—31
18	22	62	15 (1)—	142	59—73
19	15	47	12 (1)—	100	44—46
21	34	37	9 (3)—	74	91—113
22	16	36	6 (3)—	89	51—45
23	26	25	3 (2)—	71	73—83
24	14	33	10 (1)—	62	48—36
25	18	38	14 (1)—	78	51—57
26	23	43	9 (1)—	75	65—73
27	32	54	12 (1)—	96	98—94
28	17	44	13 (1)—	90	51—51
29	20	50	18 (1)—	105	51—69
30	11	20	3 (2)—	56	28—38
31	38	40	3 (2)—	100	115—113
32	42	17	3 (5)—	70	137—115
33	24	33	2 (3)—	73	93—51
34	34	36	2 (3)—	88	123—81
35	29	36	0 (2)—	124	89—85
36	24	28	2 (4)—	89	96—48
37	16	13·5	0 (3)—	65	64—32
38	21	8	1 (10)—	25	87—39
39	34	14	0 (8)—	50	130—74
40	15	29	0 (2)—	87	57—33
41	30	35	6 (1)—	78	93—87
42	27	23	1 (2)—	59	94—68
43	29	20	0 (3)—	52	100—74
44	30	9	0 (11)—	32	115—65

The first column of Table VII gives the sequence of the generations of tank rats (for short, T.Rs.) trained from the 13th to the 44th generation.

[1] The tables of figures in this report are numbered consecutively to those of similar form in the Second and Third Reports.

In this connexion I repeat the explanation that the method of procedure in this experiment was not fully worked out until the training of the 13th generation had been begun. The numerical measures recorded for the preceding 12 generations are therefore not comparable with those for the 13th and later generations.[1]

Something of the nature of the changes of procedure has been indicated in the foregoing reports. Here it may suffice to say that the procedure of the earlier stages, while conducing to the same result, namely the learning to avoid the bright gangway as consequence of shocks received on it, was very much less prolonged and thorough than that of the subsequent stages. The training effect on each of these first 13 generations must therefore be regarded as very decidedly less than that on the subsequent generations. I estimate it at 30–40 %, or roughly one-third, of the value of the latter. If this estimate is approximately correct the 13th generation of the tank rats may be regarded as equivalent (in respect of amount of preceding training of the stock) to the 5th or 6th generation of a series subjected to the full training procedure from the outset.

Certain features of Table VII may be pointed to forthwith as indications of progressive increase of facility in the mastery of the task set to each and all of the rats, with the increase of the number of generations trained in the tank:

(1) A decided though not regular falling-off in the average number of errors (=shocks received), made by the rats of the groups of successive generations, is obvious on inspection of the third column[2].

(2) A similar falling-off in both columns of the double fourth column, representing the range of error, i.e. the smallest and the largest number of errors made by any of the rats of each generation. Here the figure in brackets states the number of animals in each group which gave the same minimum number of errors; this figure becomes of importance in the lower part of the table where the better rats achieve, or approximate to, mastery of the task without errors.

(3) The double fifth column gives the results of zero-day testing. On

[1] This is an unfortunate and irremediable defect of this series. I now regret that I did not initiate the series afresh with new stock at this point. However, such a new series was later initiated, the W.C. (W.H.) series, with further improvement of method. The working out of a relatively perfect method of procedure is a much more difficult task than might be supposed and than I myself had supposed at the outset. It can be safely said that it is only possible to a psychologist of all-round experience and equipment.

[2] In correcting proof I add that 34 Tank Rats of the 46th generation have made only 3 errors per rat; this figure multiplied by 1·54 gives the corrected average error for this generation as 4·6.

zero-day (the day preceding the beginning of training with the aid of the electric shock) each rat was immersed six times in the tank, all conditions (including alternation of the bright light from one gangway to the other) being exactly as in training, except for the absence of electric shock. In the early years of this tank experiment I had relied upon this zero-day performance as a possible revealer of some transmission of the effects of training, i.e. of the induced aversion from the brightly lit gangway. Having failed to find any but faint indications of this kind, I continued the zero-day procedure as a useful preliminary to the training process. It was not until I reached the 34th generation that my interest in the zero-day performance, as a possible revelation of a Lamarckian effect, was revived by noticing that the groups of rats of the 32nd, 33rd and 34th generations had shown a decided tendency away from *B*, the bright gangway, or, at least, a marked preference for *D*, the dimly lighted gangway. The lower part of the fifth column shows that this preference (which, there is hardly room for doubt, is founded in an aversion for *B*) has been consistently and increasingly manifested throughout the fourth period of the experiment, i.e. the period since the third report, comprising the 10 generations 35–44. *This preference has thus become one of the leading evidences which I adduce in support of the Lamarckian interpretation.*

I may add here that, in view of the importance and the extreme delicacy of this part of the experiment, I have used during the fourth period the utmost care in conducting these zero-day observations. Each rat was placed alone in the tank at the normal starting point, being held by the tip of the tail and dropped in gently; while I, in order to avoid any possible influence upon his behaviour, forebore even to allow myself to watch him continuously; contenting myself as a rule with cautious peeps over the edge of the end of the tank, if he did not quickly emerge on one or other gangway. The subdued illumination of the room was carefully regulated to avoid any complicating variations thereof.

In order to have further data for estimating the significance of this zero-day preference for *D*, I cite all the corresponding figures for control groups. And there is here a complication, namely, the fact that, in the same year (1932) in which this decided zero-day preference for *D* began to be manifested, it became necessary to build new tanks to replace the old pair which had been in use during some 10 years. The pair of new tanks, although modelled closely upon the old ones, proved to offer to the rats a less difficult task than did the old ones, less difficult by roughly one-third; or, exactly (as determined by a long series of comparative trainings extended throughout many months, cp. (3), p. 217), the difficulty

of the old tanks was to that of the new ones as 1·54 to 1·00; and all figures expressing degrees of facility of any rat or groups of rats trained in the new tanks have been corrected by multiplying the number of errors made by 1·54.[1]

Now it might be suggested that the well-marked zero-day preference for *D* of the later generations is due merely to their having been trained in the new tanks. And in fact there is good evidence that it is in part, but only in small part, due to this circumstance. There is equally clear and abundant evidence that in the main it is not due in more than small part to this circumstance; this is provided by Table VIII which sets out the performances of two large groups of control rats.

Table VIII

Part I. Seven groups of controls trained in old tank.

Stock	Average no. of errors per rat	Range of errors	Zero-day choices *D*—*B*
17 W.Cs. (1926)	149 +	132—165 +	47—55
10 W.Cs. (1927)	164	90—299	28—32
22 W.Cs.$_2$ (1927)	149	66—234	57—75
20 W.Cs. (1930)	117	60—237	61—59
11 S.Cs.	188	16—282	29—37
4 D.U.Cs.	234	159—351	12—12
22 Bath controls	107 +	69—129 +	68—64
Total of 106 controls	143	—	302—334

Part II. Five groups of controls trained in new tank.

14 hooded Cs.	178	50—308	41—43
14 C.H.Cs.	125	41—191	44—40
12 N.Y.Cs.	76	32—124	39—33
14 D.U.Cs.	174	51—300	42—42
49 W.Cs. (1932)	102	14—352	147—147
Total of 103 controls	122	—	313—305
Total of 209 controls	133	14—352	615—639

One feature of Table VIII is the evidence afforded by the totals of zero-day performances, to the effect that, while in the old tanks there is manifested by 106 controls a very slight preference for *B* (the bright gangway), in the new tanks 103 controls manifested no preference for either *B* or *D*, taking the two routes in almost exactly equal numbers.

[1] Here an exception must be explained. In stating the minimum number of errors made, this correction has not been made when the minimum number was zero, 1, 2 or 3. Clearly, zero cannot be multiplied by 1·54; and it seemed that the figures 1, 2 and 3 uncorrected would give a truer picture of the performances than the awkward fractions 1·54 errors, or 3·08 errors, or 4·62.

The evidence of this table suffices, I think, to establish the point that the very marked zero-day preference for *D* shown by the later generations of the tank rat stock cannot be attributed to the change from the old to the new tanks, save in some very slight degree. But there is other evidence strongly supporting it to be found in other tables, namely, that in no series does the appearance of the zero-day preference for *D* coincide with the transition from the old to the new tanks.

Since I have no comparable records of the performance of the 1st generation and of the succeeding eleven early generations of the tank rats, I have to rely upon comparison of the performances of the later generations with the performances in the same procedure of many groups of control-rats of the same Wistar stock and of other stocks. The performances of all the control-groups dealt with in the three earlier periods figured in my Third Report ((3), Table III). Table VIII presents these same figures, supplemented by those representing performances of several new control-groups trained during the fourth period; the whole table being divided (for reasons already stated) into two parts, representing controls trained in the old and in the new tanks respectively.

The discussion of the control-groups contained in my Third Report may be supplemented by a few words here. It is obvious on inspection of Table VIII that the two parts agree closely. The group 12 N.Y.Cs. is somewhat anomalous in that these 12 rats (obtained from a New York dealer and of unknown ancestry) made by far the best performance of all the many different control-groups. It will be seen that, with the exception of this one group, no one of the 12 groups has given an average error per rat as low as 100. My previous estimate of 120 as the most probable figure representing the facility of the 1st generation of tank rats may still hold as the lowest reasonable estimate. The fact that the 14 hooded rats gave so high an average of errors as 178 is of some little interest. I had expected them to show decided superiority to the albino rats: first, by reason of the presumably better vision of their pigmented eyes; secondly, because I have assumed, as a high probability, that the albino rats (which must be supposed to have lived for many hundreds of generations confined in small cages in the main) are somewhat degenerate as compared with the wild stock of *Mus Norvegicus*; the albino stock must have been shielded in large measure both from all natural selection tending to maintain its qualities and from activity which might maintain or improve its qualities by Lamarckian transmission. I regret that I do not know the derivation of this hooded stock of which the rat colony of my department had a considerable quantity.

Table VII continues to show in its later part the very considerable ups and downs of the figures representing the average number of errors for the rats of the successive generations.[1] I have thought it worth while to smooth out the curve by amalgamating the figures for groups of generations. There are 31 generations represented in Table VII. In Table IX these are divided into four groups, the figures for each group representing the amalgamations of the figures for the first, second, third, and last series of 8 generations.[2]

Table IX

Generations of T.Rs. in groups of 8	No. of rats	Average no. of errors per rat	Range of errors	Zero-day choices *D*—*B*
13th to 21st generations	130	56+	9—147	374—406
22nd to 29th ,,	166	41	3—105	488—508
30th to 37th ,,	218	29	0 (5)—124	745—562
37th to 44th ,,	202	20	0 (19)—87	740—472

Table IX serves to bring out clearly the marked falling off in the number of errors made in the three columns representing, respectively, the average number of errors per rat, the number of errors of the quickest rat, and the number of errors of the slowest rat. In the last double column, I remind the reader, the point of interest is the difference between the figures under *D* and *B*, the former representing the number of times the dim gangway was taken on zero-day, the latter the number of times the bright gangway was taken. Up to the 30th generation there is slight excess of *B*'s; after that an increasing and great predominance of *D*'s.

In the earlier reports I have described certain peculiarities of behaviour in the tank on the part of tank rats of the later generations. I may say in general terms that in the 10 generations of the fourth period these peculiarities, consisting in evidences of extreme caution and of aversion from *B*, manifested before receiving any shocks or after very few shocks, have become further accentuated ((3), p. 219). One such peculiarity deserves special mention, since it has become very frequent and pronounced during the fourth period: namely, the making momentary contact with the wire-mesh on *B* followed by immediate and violent springing back from *B*, after receiving through nose, paw, or whisker, some degree of

[1] I remark that Dr Crew's comparable table (to be discussed in Part II) shows equally numerous and more extreme variations of this figure from generation to generation. The numbers of animals he dealt with in each generation were considerably larger than in my series.

[2] This gives a harmless overlapping in that the 37th generation is represented in both third and fourth amalgamated series.

shock from the wire. This springing back is usually followed by a swift swimming by the most direct route to *D* and emergence by that route[1]. Closely allied with this 'contact' behaviour and manifested in the main by the same rats was the following behaviour. After emerging from the water by *D*, the rat rests quietly on the platform, *P*, showing perhaps a quiet alertness in his general bearing; then when the switch is thrown over, bringing the bright light to the gangway above which he is resting (but without the light shining directly upon him and without any of the shock-current reaching him) the rat makes a sudden dash down the gangway, leaps into the water, swims round to the other gangway (now *D*) and emerges thereby. I cannot assert roundly that this behaviour is absolutely peculiar to the last few generations; but I cannot remember seeing it manifested by rats of any other groups.

Yet another significant peculiarity of behaviour has become very frequent in the last few generations of tank rats. In the earlier generations and in all the control groups, it has been the general rule that after a longer or shorter period of delay or loitering at or about the starting point, each rat sets out and swims vigorously and without further loitering or hesitation to either *B* or *D*, and promptly climbs out or rather, after some little experience of the shock, rushes out at maximum speed. Among the tank rats of the last few generations, not only have 'con-

[1] Here I may explain a feature of my procedure of some slight importance, one which Dr Crew has imperfectly apprehended and somewhat misrepresented in his article. In my old tanks each of the glass wire-covered gangways was so placed that its lower edge was just clear of the surface of the water. In order to prevent any rat from swimming under this edge, a vertical partition (at first of glass, later of wood) was submerged immediately below this lower edge of the gangway, its upper edge being just below the surface of the water. The rats could and often did seize this upper edge using it as a step to the gangway. In my new tanks the ground-glass gangways slope down to the bottom of the tank; the wire mesh covers only the upper part of the glass, its lower edge being about 5 mm. above the water. The place of the upper edge of the vertical partition which served as a step to the emerging rat is taken by a broad rubber band (1 cm.) stretched tautly round the glass plate, its upper edge flush with the surface of the water. Without this step, the rats are liable to scrabble helplessly at the glass surface; although if they approach with a rush they may reach the wire without making use of the rubber band. One interesting variety of the contact behaviour, not very uncommon in the latest generations, is that the rat slowly and cautiously approaches *B*, catches hold of the rubber band, and then, slowly advancing his nose towards the wire, makes a very slight contact, receives a very slight shock, and withdraws his head without losing hold on the rubber band; and perhaps repeats this several times before letting go and swimming round to *D*. The question arose whether such 'contacts' were to be reckoned as errors in the many cases in which the contact was followed by retreat from *B* and emergence by way of *D*. I adopted the convention of not reckoning such a contact as an error. If each such contact had been counted as an error, the number of cases of zero-error would have been considerably reduced.

tacts' been frequent; but they seem to be advancing to a further stage of caution which tends to obviate the contacts; namely, it has frequently occurred that a rat will swim towards *B*, and then, instead of rushing at it or, as in 'contact' behaviour, cautiously approaching *B*, he stops some few inches before reaching *B*, turns round and swims as directly as possible to *D* and forthwith emerges by way of *D*. This is called 'turning back'.

It seems worth while to give a more complete picture of the behaviour of one batch of 9 tank rats of the 44th generation, the best performers of any group hitherto trained. Zero-day performance is 38 *D*–16 *B*, but all 9 rats emerge at least once by *B*:

1st day of training: 9 rats make 7 errors in all, but they also make 10 contacts. Five rats make no error on this or subsequent days, i.e. after zero-day they never emerge by *B*.
2nd day: 9 errors are made, and 2 contacts.
3rd day: 5 errors, few contacts and much turning back from *B*.
4th day: 4 errors
5th day: 3 errors } Still few contacts and much turning back.
6th day: 0 errors, all runs perfect.
7th day: Ditto.

This group of 9 tank rats thus made a total of 28 actual errors, and about 18 contacts. Five made no errors and the average number of errors per rat was 3·1.

(4) *The adversely and the favourably selected series of generations trained in the tank*

In order to rule out the possibility of some unwitting favourable selection, I initiated in the year 1928 a new experiment differing from the tank experiment dealt with above only in that, instead of random rejection of the superfluous progeny of each generation, I rejected from each batch of rats trained (generally 6–10 rats, either of 1 litter or made up from 2 or 3 litters) the half which made the better performances and kept only the slower half of the batch for breeding the next generation. (In the case of uneven number, the odd rat was usually kept for breeding —thus of a batch of 7, the best three performers were discarded; the 4th, 5th, 6th, and 7th in order of merit were kept for breeding, cp. (2), p. 205.) The training began with the first generation of rats newly obtained from the Wistar Institute of the same stock as the tank rats; the adverse selection began with the 3rd generation. Table IV ((3), p. 224) gives the performance of this series of generations from the 1st to the 16th. Table

X of the present Report continues this series to the 24th generation and adds the zero-day performances.[1]

Table X is constructed according to the same rules as Table VII. Before making a few comments upon it, I present the similar Table XI which presents the corresponding data for the favourably selected series

Table X

Generations of W.Cs. (W.H.)	No. of rats	Average no. of errors per rat	Range of errors	Zero-day choices *D*—*B*
3	5	215	204 (2)—229	15—15
4	18	166	75 (2)—241	53—55
5	11	134	81 (3)—196	32—34
6	19	119	63 (1)—194	60—54
7	15	115	54 (1)—150	41—49
8	9	147	69 (1)—231	31—23
9	14	139	51 (1)—228	39—54
10	7	114	69 (1)—192	21—21
11	15	129	81 (1)—198	39—51
12	21	81	24 (1)—165	53—73
13	13	59	20 (1)—120	37—41
14	41	58	14 (1)—144	122—130
15	36	70	23 (1)—148	111—105
16	65	43	3 (13)—129	250—140
17	61	43	2 (1)—140	171—195
18	33	45·5	1 (2)—106	124—74
19	20	38	11 (1)—74	50—70
20	36	47	1 (2)—109	131—85
21	26	56	14 (1)—131	83—73
22	22	34	3 (2)—87	85—47
23	34	52	3 (1)—76	83—121
24	29	27	1 (3)—116	107—67

W.C. ($B.H._2$). This series (the first attempt at which came to naught by reason of the sterility of the stock concerned)[2] derives from the 14th generation of the W.C. (W.H.) stock. Instead of discarding the 21 better

[1] Some explanation is due in regard to the last generation listed in Table X. The 29 rats of this generation include a litter of 11 rats which suffered a fatal accident. (My assistant, with excess of zeal, and without my advice, undertook to dip all rats of my stock in disinfectant as a precaution against skin parasites. There were numerous fatal casualties; but the incidence was heaviest on this my most valued litter which was exterminated.) At the time of the accident, 7 of the 11 rats of this litter had completed their training; 4 were still making errors at the normal average rate of 3 a day. The 7 had made on the average 11 errors per rat. Clearly it is not legitimate to include these 7 in the record without the other 4 whose training was not complete. I have therefore assigned to these 4 the same number of errors as were made by the slowest 4 of the batch of 7 rats of the same generation, whose training had just been completed. It is almost certain that this number is higher than these 4 rats would actually have made. Without these 11 rats the remaining 18 rats gave an average error of 31, which still surpasses all previous records. In correcting proof I add that 27 rats of the 25th generation made 29·5 errors per rat; 42 rats of the 26th generation made 35·6 errors per rat.

[2] Cp. (3) Table VI, p. 232.

performers of this group, W.C. 14 (W.H.), of which there were 42 members, I interbred these 21 rats and designated their progeny W.C. 15 ($B.H._2$). The series thus initiated has been perpetuated on the same principle, namely, of breeding from the better half of the members trained in each generation and discarding the worse half.[1] The purpose of this experiment is to throw light on the question—What degree of improvement of the facility trained and tested by the tank procedure can be effected by consistently favourable selection in this closely inbred stock?

Table XI

Generations of W.C. ($B.H._2$)	No. of rats	Average no. of errors per rat	Range of errors	Zero-day choices *D*—*B*
15	26	50	3 (1)—146	57—63
16	34	39	3 (1)—80	127—77
17	20	44	3 (1)—124	50—70
18	27	50	9 (1)—107	91—71
19	18	51	9 (1)—131	51—57
20	12	50	27 (1)—110	41—31
21	23	28	1 (1)—76	57—81
22	23	57	13 (1)—112	85—53
23	29	49	3 (1)—111	108—66
24	20	39	3 (1)—99	72—48

Comparison of Table XI with Table X shows very little, if any, superiority of the favourably selected generations to the corresponding adversely selected generations. This series thus confirms on a larger scale the experiment W.C. ($B.H._1$) ((3), p. 232). Table X shows progressive improvement in all columns closely comparable to that of Table VII; and, just as the adverse selection seems hardly to have retarded the improvement, in comparison with the strictly comparable W.C. ($B.H._2$) series, so comparison of generations 13–24 in the two Tables VII and X (the only strictly comparable parts of the two tables) shows only slightly more rapid advance in the tank rat series. In making this last comparison, it must be remembered that the training of the first twelve tank rat generations was much less prolonged and thorough than that of the later part of the series and of the whole W.H. series: a fact which perhaps may account for the somewhat more advanced state of the 13th generation of the W.H. series as compared with the same generation of tank rat series. The same fact perhaps accounts also for the earlier appearance in both W.H. and B.H. series (remarkably parallel to one another in this respect)

[1] It should be said that the process of discarding one-half of the rats trained has proved sufficient to keep down these stocks to manageable size, so that I have trained every rat of each generation except those of the discarded half. This applies to both stocks W.C. (W.H.) and W.C. ($B.H._2$).

of the predominance of *D* choices on zero-days. In both of these series a marked predominance first appears in the 16th generation; whereas in the tank rat series a similar predominance does not appear until the 32nd generation. The difference may seem too great to be explained in this way. But it is to be noted that in the tank rat series the strong predominance of *D*, having appeared in the 32nd generation, is thereafter steadily maintained to the end of the series (13 generations) and approximates throughout to a 2 to 1 ratio; while, on the other hand, in both the W.H. and the B.H. series, the predominance of *D* is neither so great, nor so regularly manifested. The last point may be illustrated by amalgamating the figures for the last 8 generations of tank rats, the last 7 of W.H. series, and the last 9 of B.H. series, which gives approximately equal numbers of rats from each series: we then have the small Table XII.

Table XII

	Zero-day choices *D*—*B*
202 T.Rs.	740—472
200 W.Hs.	663—537
206 B.Hs.	682—554

As regards the various signs of extreme caution in the tank, so strongly manifested by the later tank rat generations, the later generations of the W.H. and B.H. series show similar signs, but the signs are less frequent and less pronounced.

(5) *Further cross-breeding experiments*

In the Second Report ((2), p. 208) is described the experiment of crossing three very slow control females with tank rat males of the 21st generation, and training the F_1 and F_2 generations of progeny, both of which made an average number of errors per rat about half-way between the number made by the mother and the father respectively.

I have repeated this experiment several times and always with similar results. I cite only two of these. In 1935 I crossed females of the stock C.H.C. (cp. Table VIII) whose average error per rat was 125, with males of the 39th generation tank rat series, whose average error per rat was 14. Of the 16 rats taken at random from the progeny and trained, the average of errors per rat was 44—range 1 (2)–102, zero-day choices 49—47. The single error made by two of the rats is a suggestion of the all-or-none and segregating transmission which is called Mendelian.

The other case is of particular interest because, unlike all the other cross-breeding experiments, it gives still clearer indication of the same kind. Of the progeny of the hooded rats mentioned in Table VIII I took at random 6 females and trained them. They made an average of 138 errors per rat (their parents had made 178); their range was from 33 to 277. These 6 were then mated with 1 male tank rat of the 43rd generation, one of those which had mastered the task without making a single error (he may have made one or two 'contacts'). Of the 6 females, 4 proved fertile and the individual performance of each one was on record. The 4 litters were trained in the usual way, beginning about the end of the 4th week; the results of training are stated in Table XIII.

Table XIII

Stock	Average no. of errors per rat	Range of errors	Zero-day choices
14 hooded rats	178	50—308	41—43
6 female hooded rats progeny of 14 hooded rats above	138	50—277	15—21
4 of the 6 females above which produced the 31 cross-bred progeny below	159	50—271	—
10 Rs. progeny of hooded ♀ of 90 errors and of T.R. ♂ of 0 errors	32	0 (1)—77	34—26
6 Rs. progeny of hooded ♀ of 111 errors and of T.R. ♂ of 0 errors	16	0 (2)—70	19—17
10 Rs. progeny of hooded ♀ of 130 errors and of T.R. ♂ of 0 errors	33	14—56	32—28
5 Rs. progeny of hooded ♀ of 306 errors and of T.R. ♂ of 0 errors	80	13—79	15—15
Total of 31 cross-breds	37	0 (3)—79	100—86

The figures of Table XIII seem to show very clearly the transmission of facility in the tank-learning through the father; his influence seems to be manifested in all three columns of the table. The average error per rat of the cross-bred progeny is, as in earlier experiments of similar kind, about midway between the average error of the four mothers and the number of errors of the father. But while no one of the mothers (or of the whole group of 14 hooded rats) made fewer than 50 errors, 3 of the cross-breeds made no errors, and others of them made very few. It is hoped to carry on the training of this series through several more generations.

PART II. REPLY TO SOME CRITICISMS OF THE FIRST, SECOND AND THIRD REPORTS

(1) *Dr Crew's criticism*

The most weighty criticism I have to meet is that of Dr F. A. E. Crew, not only by reason of his high competence, but also because his remarks are made in the light of his personal conduct of a very similar experiment, extending through some five years (4).

Before considering his critical remarks, it is in order to ask whether the negative result of Dr Crew's experiment does not in itself suffice to render unnecessary any further consideration of my results. This is the view which, I imagine, has been taken by wellnigh all biologists who have read Dr Crew's article, or even merely become acquainted with the fact that he has (in a sense) repeated my experiment and reported that in his hands the experiment has yielded no indications of Lamarckian transmission. If nothing could be said to explain the very great difference in this respect between the results of his experiment and of mine, I should have to accept his results and his verdict as reducing the weight of mine to negligible proportions.

(2) *Crew's use of an easier task*

But, fortunately for the validity and bearing of my results, there is something to be said by way of explanation of the difference. The main point of such explanation is that Crew's procedure, although it was so similar to mine that it can be described in the same general terms (as on p. 322 of Part I of this Report), was in so far different from mine *as to provide for the rats a very much easier task.* In the years during which I was devising and improving the tank procedure, an aim kept prominently in view was to make the task as difficult as possible (i.e. as near the upper limit of the rats' adaptive power or intelligence, as would test or evoke that adaptive power to the utmost), without going beyond the point at which the task of adaptation would be impossible to the less intelligent animals. The fulfilment of this condition of maximum difficulty seemed very desirable, perhaps indispensable, on two grounds: first, the more difficult the task, the more widely differentiated in respect of their facility in mastering the task would be the rats trained by the procedure; secondly, the more difficult the task, the more repetitions of the training process would be necessary for the mastering of the task; and it was a fair assumption that, if effects of training are in any degree transmitted, the

more repetitions are involved in the learning process, the 'deeper' will be the effect upon the individual learner and the more probable some transmission of that effect.[1]

These initial assumptions and the procedure which they led me to adopt have been justified by the course of the experiment. The task presented by the alternating procedure in my tank has proved to be within the scope of every one of many hundreds of rats subjected to the training; and yet it has been of such difficulty that many rats have required 200 repetitions of the shock, and some others more than 300 repetitions, before they have learnt to avoid it.[2] And my 209 control rats made on the average somewhat more than 130 errors per rat. Further, it has become very clear that an essential step of the process of learning is normally (at least, in all rats except, possibly, the few that master the task very quickly, such as those which occur in the last few of my T.R. series) the gradual acquisition of an increasingly strong aversion from *B*, an aversion which seems to be strengthened with every repetition of the reception of the shock.

The behaviour of the rats in the tank affords *abundant evidence of the gradual growth of this aversion,* largely in the form of increasing caution and reluctance in approaching *B*. Some of these manifestations have been mentioned in the Second and Third Reports. I have become increasingly interested in them as it has seemed to become clearer that just this slowly acquired aversion from *B* is the main feature (perhaps the essence) of that which seems to be transmitted in Lamarckian fashion. The whole phenomenon of a prolonged preliminary period of the learning process preceding the achievement of discrimination and corresponding adaptation of behaviour is one of great interest in connexion with the general theory of the learning process in men and animals. Working in my laboratory, Dr T. L. McCulloch and Dr J. G. Pratt(5) have made an elaborate study of it by highly ingenious experiments; here the main interest of their study is the demonstration of the very real and important nature of this 'pre-solution period', as they have proposed to name it. Here I will merely cite one very convincing kind of evidence of its reality presented by many of my rats. As said before a considerable proportion of all rats form, in course of training in the tank, a 'side-habit'; i.e. the rat acquires the habit of leaving the middle chamber into which he has been placed always by the same one of the two exits (right or left) and emerging from the water by the corresponding gangway. It happens in very many such cases that the rat, having acquired this side-habit (say, to make exit by turning to the left on every immersion), subsequently acquires an increasing aversion from *B* and manifests it in the following fashion (of great interest from the point of view of psychological theory). Before each immersion the light and the shocking cur-

[1] Cp. Dr H. S. Jennings on the time factor in Part III of this paper.

[2] It is to be remembered that 300 repetitions of the shock to any rat mean roughly 600 immersions of that individual; or a course of training extending through 100 days.

rent are switched in routine fashion from right to left or from left to right. At first there is no obvious difference in the behaviour of the rat at and about the starting point, *O*, according to the position of the light; he commonly, after a very brief hesitation, swims swiftly to the left and so out, manifesting perhaps some expectation of the shock by his hurried pace or rushing as he approaches *B*; and also, in some cases, as he approaches *D*. But after some further days' training, he begins to loiter at or near *O*, when the light (and the shock) are on the left; but continues, when the light is on the right, to swim out to the left with a minimum of hesitation or delay. Prof. Agar (see Part III below) reports similar behaviour. It is clear that this difference of behaviour at *O* must be based upon perception of the light reflected from the end-wall of the tank on the left side; for that is the only sensory cue which can serve to cause the loitering at *O*. And, as the training proceeds, the increase of strength of his aversion from *B* is manifested by increase of the period of loitering at or about *O*, before he will start on his journey; and also, in some cases, by turning back after starting and swimming some smaller or larger part of the habitual route to the left-hand gangway, followed by a turning-back which may be repeated many times, a number increasing with the further repetition of shocks received on *B*. (Agar remarks that none of his rats has shown this turning-back behaviour. It is common in my later T.R. generations and is one of the peculiarities which distinguish them from control groups.) It seems quite clear that the rat exhibiting such behaviour becomes the seat of a conflict between two opposed impulses, whenever he is placed in the water at *O* while *B* is on the left side, a conflict which is not evoked when *B* is on the right side; a conflict between on the one hand the impulse to escape from the water by swimming to the gangway on the left (a directed impulse evoked at once by immersion at *O*) and, on the other hand, an impulse of aversion or avoidance of *B*, which is evoked in *R*, not only by the sight of *B*, but also, while *R* is still at *O*, by the light reflected from the leftward part of the end-wall of the tank, which light serves in some sense as a warning that the shock may be expected on that side. All this seems, no doubt, extravagantly complicated to pure behaviourists and reflexologists who do not believe that a rat has any mental life involving anticipations or foresights which govern his behaviour. But just that is one of the inevitable difficulties of this problem. Events in the animal are not so simple as the reflexologists and some geneticists seem to believe.

If then Crew's procedure differs from mine in that it does not even distantly approach the *desideratum* of *maximum difficulty*, but rather offers to the rats a very much easier task requiring only a fraction of the repetitions of the shock required in my procedure, it follows that Crew's procedure, in spite of its general resemblance to mine and its superiorities in minor respects, lacks one of its most important features, and is not well suited to reveal any possibility of Lamarckian transmission which may be present in the species. Let us then survey the evidence of this great flaw in Crew's procedure.

Abundant evidence is provided by comparison of the numbers of errors made by his 18 control groups with those made by my 12 control

groups. Crew trained 18 control groups each containing 50 or (in a few cases) more than 50 rats.

(1) The average error per rat of all these hundreds of controls was 40·50; whereas the average error per rat of my 209 controls was 130 or a little higher.

(2) Crew's best control group made 18·08 errors per rat; my best control group (the 12 N.Y.C. of Table VIII), which was much superior to all others, made 76 errors per rat; no other making fewer than 102 errors.

(3) The worst of Crew's control groups made 72·46 errors per rat; the worst of my control groups made 234 errors per rat.

(4) Ten of Crew's controls made no errors at any time and his table of ranges shows that beside the 6 groups which ranged as low as 0 errors, 2 groups ranged down to 1 error, 5 ranged down to 2 errors, 3 to 4, 1 to 5, and 1 (the slowest group in this respect) to 8 errors; while of my 12 groups, 1 only ranges as low as 14 errors; 4 only range below 50 errors, and 2 do not range below 100 errors.

It is the figures of the last paragraph which bring home to one most clearly this defect of Crew's procedure, the defect namely that it presents too easy a task and requires too few repetitions of the shock. For they make it obvious that many of his rats received only very few shocks. Indeed Crew tells us that of his 18 generations successively trained by this tank procedure (what he calls his experimental generations), 29 rats actually went through without a single shock. These must therefore be dismissed as having received no training at all; while it is clear that of the total number of animals in these 18 groups (1449) *a very considerable proportion, over and above these* 29 *entirely untrained rats, received a very slight training only, involving very few shocks to each rat.*

It is obvious that in so far as the purpose of the procedure is to bring to light any transmitted effect of training in successive generations, a form of procedure which omits all training from some members of each generation and gives to many others only a very slight training is *ipso facto* disqualified to throw light upon the problem in hand.[1]

(3) *The variability of errors in the control groups*

Inspection of Crew's Table II (which he has been so kind as to draw up, for the sake of ease of comparison, in the imperfect form adopted by myself) reveals another and more obscure and perhaps even more grave

[1] What exactly is the ground of the greater difficulty presented to the rats by my procedure as compared with Crew's I am not prepared to say; but two factors are in my view of chief importance: (1) the different dimensions and proportions of the parts of the tanks;

defect of his method. I refer to the very large differences in the numbers representing the average score, i.e. the average number of errors per rat, in the 18 control groups. The figures range from 18·08 to 72·46, the latter being almost exactly 4 times the former. Crew may retort that the same defect is chargeable to my procedure, because in my Table VII the figures representing the average error per rat for the 31 groups of rats range from 8 to 80. But to do this would be to beg the question at issue between us. Between the group which gave the figure 80 and that which gave the figure 8 intervened 22 generations of rats all intensively trained in my tank. And it is my contention that the difference between 8 and 80 (or a large part of that difference) represents improvement of facility due to Lamarckian transmission. It is Crew's endeavour to explain away the difference. The difference between Crew's best control-group and his worst is slightly more than 54 (i.e. 72·46 less 18·08): if my Table VII showed any adjoining groups differing in respect to average error per rat as widely as 54, my procedure would be open to the same charges as Crew's, namely, that it gives a very wide margin of experimental error, a margin so wide as to be quite capable of masking completely any small evidences of Lamarckian transmission which a more accurately measuring procedure would have revealed. It must be remembered that there is no assignable ground why any considerable differences in respect of average error per rat should appear between any of the 18 control groups of Crew's series. In Crew's Table II the variations of the figure expressing the average number of errors seem to be purely random; e.g. about the middle of the table, the figure makes the following violent jumps: 59—21—32—69—25—23—34—52—72—47. These large differences without assignable cause other than experimental error (or latitude or looseness of procedure employed in measurement) are the more significant in Crew's Table II in view of the fact that no one of the groups has less than 50 members; the numbers in the groups being thus sufficient to smooth out very largely differences which in smaller groups might properly be attributed to chance inclusion in one group of a number of unusually

(2) the very different illuminations in his tank and in mine. The latter seems to me the more important. In Crew's tank the rat, turning from the middle passage towards *B*, enters at once a blaze of light from a lamp immediately above him; whereas in my tank, he enters a long passage at the far end of which (24 in. distant) is the glow lamp, separated from him by the ground-glass gangway (closely covered with wire-mesh and therefore very obstructive to the passage of the light) sloping to the surface of the water at a point 12 in. or more from the rat. Instead of entering a blaze of light, then, he enters a passage at the far end of which is an oblong panel lit up, but not very brightly, by the lamp 12 in. behind it; a very different situation. One can readily understand that the former is the more likely to evoke an immediate retreat or unconditioned aversion.

capable individuals and, in another group, of one or more exceptionally slow learners. No such wide differences between the groups of immediately successive generations as 43, 36, and 38 (which appear in Crew's table) occur in my Table VII. In that table the widest differences between successive groups are 30 and 27; and in both cases one of the two groups concerned is so small (namely 11 members) as to make it very probable that the figure does not approximate to a true average for the group. If these two instances be put on one side on that ground, the three largest figures expressing differences between adjoining groups (to parallel the three corresponding figures from Crew's table 43, 36, 38) are 23, 22, 15.

From the point of view of comparing the objective validity of the figures expressing average number of errors obtained by Crew's procedure with those obtained by my procedure, it would be more satisfactory to compare my control groups with Crew's, but for the fact that most of my control groups are too small for this purpose. However, of my 12 control groups 5 contain each not less than 20 rats, and within these 5 groups the figures representing the average error per rat for the 5 groups range from 102 to 149. This I submit compares very favourably with the range from 18 to 72 of Crew's 18 control groups; especially in view of the fact that of these 5 largest of my groups one only approximates in number of its members to Crew's groups; and of the further fact that the rats of all Crew's 18 control groups are very much more closely allied by blood than the members of my several control groups. In Crew's control series, each group consisted of the progeny of litter-mates of the members of the preceding group and, like the rats of his 'experimental series' also, were all derived from two pairs of rats, and were thus extremely closely inbred and homogeneous; whereas my five largest control groups, although all derived ultimately from the same Wistar Institute stock, were the progeny of larger groups obtained from that stock in five different years. The members of Crew's successive control groups were related as nephew to uncle; those of my control groups as distant cousins 'many times removed'. For these two reasons (larger numbers in each group, and closer kinship), Crew's table should, if his procedure had been equally objective and accurate, show decidedly smaller differences between groups than does my Table VIII; rather than the considerably larger differences actually shown.

In support of both the foregoing reflections upon Crew's procedure, I cite here a table which I furnished some years ago to Dr Crew and which he has published in his article as Table V ((4), p. 93). In 1931 Dr Crew

very kindly gave me 12 rats of the 5th generation of his 'experimental stock', which I transported from Edinburgh to North Carolina via Montreal. One of these succumbed to an accident; 11 were trained as Scottish controls of the 1st generation ($S.C._1$), and groups of their descendants (selected at random, but evenly, from all litters born) were trained by Rhine and myself in our tanks with the results shown in Table XIV.[1]

Table XIV

Stock	No. of rats	Average no. of errors per rat	Range of errors	Zero-day choices *D*—*B*
$S.Cs._1$	11	188	16—282	29—37
$S.Cs._2$	20	87	15—315	56—64
$S.Cs._3$	14	63	28—141	46—38
$S.Cs._4$	16	59	14—174	42—54
$S.Cs._5$	12	48	18—100	36—36
$S.Cs._6$	19	61	14—203	69—45
$S.Cs._7$	21	46	4—93	82—44

The most striking feature of this table is the very large drop in the average-error figure from the 1st to the 2nd generation. Our figure, 188 for the average error of this 1st generation, contrasts sharply with Crew's corresponding figure, namely 30·27. It is very possible that our figure is unduly high; partly, though improbably, owing to the fact that the training of the rats began within a week of the termination of their long journey (but they had been well cared for on the journey, and manifested no signs of having suffered in health).[2] It is more probably due to the fact that the number of rats in our group, namely 11, is too small to make it a true sample of the stock. The number in the group is unfortunately small; but there is no mistake about their poor performance. Rhine trained 5 rats of the group in entire independence from me, and found an average error of 149 for them, with range from 16 to 265, while under my training the other 6 rats gave 220 errors per rat, ranging from 140 to 282.

I may say that there was manifested by these 6 rats no sign of any rat mastering the task before the 48th day of training, on which day one rat ceased to make errors; a second rat achieved the task on the 51st day; the third on the 56th day; the fourth seemed to be on the point of mastery about the same time, but continued to make errors up to the 70th day; the fifth finished on the 90th day, the sixth on the 97th day. These last

[1] I have here to apologize to Dr Crew for the fact that the version of this table which I sent him and which is reproduced in his article had not been subjected to the laborious and minute revision of all figures and arithmetic which is necessary in all cases before publication and has been slightly emended in some details in the light of such revision.

[2] Cp. in this connexion some evidence on the influence of vigour on performance discussed in Part III of this article.

two had formed an obstinate side-habit, which is highly unfavourable to solving of the problem; but this was not the explanation of the long training required by the others—notably the fourth rat continued throughout to emerge on both sides without marked preference for either, yet still persisted in making errors.[1]

The rat which made the largest number of errors, namely 282, does not then stand far apart from his fellows as one much slower than the rest; two others of my group of 6 came near him, and also one of Rhine's 5 which made 264 errors.

The slow learning of this group is all the more striking by reason of the fact that I, being aware of the low figures representing the average error for Crew's first few generations, was careful to avoid any possible ground for slow learning in the one somewhat variable factor of importance; namely, the strength of the shock received by the rats. In order to ensure that this was not less than in the case of my other rats, I deliberately kept the strength of the induced current somewhat above the normal level, and have a note to the effect that tetanization was, accordingly, more frequent than normally; i.e. more than 50% of all cases.

Apart from the large drop in errors from the 1st to the 2nd generation, the average-error figures of this table show only small differences between successive generations, the largest being 15; which again compares very advantageously with the large differences of Crew's Table II.

The degree of improvement of this group within these 7 generations of training implied by the figures seems almost too great to be true, bringing as it does the 7th generation to the level of average error first reached by my 17th generation of the T.R. series. But it is to be remembered that my Scottish control series had been trained through 4 generations by Crew, so that my 7th generation S.C. is really the 11th generation of the series to be trained. On the other hand, as I have pointed out on an earlier page (p. 324), the training of the first 12 generations of my T.R. series was very much less thorough than that of all later generations, so that the 17th generation in my Table VII may be regarded as not far from the equivalent of $S.C._7$ in respect of the amount of training received by the preceding generations. Again, it is to be remembered that, for one reason or another, the group $S.C._1$ were slow to learn and consequently suffered a very long and severe training process; and this

[1] It would be a mistake to suppose that the side-habit is the ground of all prolonged persistence in error-making. This rat is one of many which I have noted as long persisting in errors, although taking both exits freely. I mention this because Crew seems to suggest the contrary.

fact may be in part the ground of the very large step of improvement from $S.C._1$ to $S.C._2$.

It is true that the amount of improvement represented by the 7 generations of Table XIV seems incredibly large if we simply deduct the last figure (46) from the first (of the average-error column) and state the difference (142) as the measure of the improvement. But it remains probable that, for one reason or another, the first figure is unduly high and the last unduly low. That some superiority of the group $S.C._7$ to $S.C._1$ (as indicated by their much smaller average error) was real and objective is borne out by the figures of the range column: here, in both columns, $S.C._7$ shows its marked superiority, not only to the small and questionable group $S.C._1$, but also to the larger group $S.C._2$.

Surprisingly, there appears also in the groups $S.C._6$ and $S.C._7$ the predominance of *D* choices on zero-days which in the T.R. series and the W.H. and B.H. series appears only after many more generations. The predominance appears too large to be attributable to mere chance. Fortunately in the case of the group $S.C._7$, the predominance was so marked as to draw my attention at once; and being naturally sceptical about it, feeling, like Crew, that the 6 trials of each rat on zero-day only are somewhat unsubstantial evidence in the case of a group standing alone (or almost alone—$S.C._7$ is supported in this respect by $S.C._6$), I subjected the rats of this group to 6 shockless trials on 3 successive days.[1] The choices of these three zero-days, on being averaged, still show a marked, though somewhat smaller, predominance of *D* choices, namely 78 *D* to 48 *B*. I feel apologetic and embarrassed about these particular figures, and would rather they had been more closely in accordance with the corresponding figures of the other tables; but there they are. I now regret very much that I did not carry the training of this stock through a larger number of generations, and can only advance my standard excuse, lack of time and energy.

At this stage I may further interrupt the flow of the argument by adducing some other evidences of appreciable improvement in facility resulting from the training of the first few generations of several series; instances which standing alone would be negligible, but which, taken together and in conjunction with the series of Scottish controls, amount,

[1] This I have done in a number of other cases in which the predominance of *D* choices was very marked. The additional shockless trials cannot be supposed to have any appreciable effect on the course of learning. But it is noteworthy that, naturally enough I think, the predominance of *D*, if very marked on first day, is generally somewhat less on second and third days.

I submit, to substantial evidence of such improvement, and serve thus to support the startling figures of Table XIV. Three such series are brought together in Table XV. The series of 1927 Wistar controls has appeared in Table I ((2), p. 205). The improvement through the 4 generations trained is steady and considerable in respect of average error and of both range columns. The same is true of the series derived from the State University at Chapel Hill, except that the very marked drop in the first range column is not maintained in the 4th generation. The 2nd trained generation of hooded controls shows similar improvement.

Table XV

Stock	Average no. of errors per rat	Range of errors	Zero-day choices *D*—*B*
10 W.Cs.$_1$ (1927)	164	90 (1)—299	28—32
22 W.Cs.$_2$	149	66 (3)—234	57—75
37 W.Cs.$_3$	144	45 (2)—229	93—129
10 W.Cs.$_4$	114	42 (1)—170	27—33
Chapel Hill controls			
14 C.H.Cs.$_1$	125	41 (1)—191	44—40
13 C.H.Cs.$_2$	87	6 (1)—180	37—41
14 C.H.Cs.$_3$	86	3 (2)—171	45—39
7 C.H.Cs.$_4$	81	46 (1)—147	23—19
Hooded controls			
8 H.Cs.$_1$	207	110 (1)—308	26—22
6 H.Cs.$_2$	138	50 (1)—277	15—21

These three short series, together with the Scottish controls series, are all the records I have of the first few generations trained successively and intensively by the standard method. Without exception they all show well-marked improvement in respect of average errors per rat and also in respect of both figures of the range column. The Table XV thus gives some support to Table XIV, which, standing alone, might seem anomalous. Table XIV, thus supported, affords substantial evidence of improvement of facility in 7 generations of trained rats, and also supports the very strong evidence of Table VIII indicating that the task presented to the rats is very much easier for them in Crew's procedure than in mine.

Crew's procedure, then, is vastly superior to mine in respect of the numbers of animals of each generation trained, and in that he recorded separately the performance of each rat and also its pedigree throughout the whole of the 18 generations dealt with[1]; but in two more essential

[1] Since in the Third Report it was mentioned that Dr Rhine was initiating a new series in which the individual records and pedigrees were to be kept throughout, I must here explain that this venture came to naught by reason of two circumstances: (1) the new stock was provokingly infertile; (2) Rhine was becoming so immersed in his parapsychological experiments as to find this procedure made an impossibly large demand on his time.

ways it was seriously defective, namely (1) the tank presented was not sufficiently difficult to insure to all the rats the long intensive training with many repeated experiences of the shock which seems desirable and is perhaps indispensable for producing a sufficiently 'deep' impression on the individual for appreciable transmission;[1] (2) so loose in its measurement of facility as to allow very large chance variations of the figures representing degrees of facility, chance variations so large as to mask variations expressive of considerable improvement of facility.

I will add that Crew's experimental series, according to the evidence of the figures of his Table II, is not entirely without indications of improvement of facility in the course of the 18 generations of training. The score of errors for the median rat of the last, or 18th, generation is the lowest of the column; the figure for the average error of the same generation is the lowest but one; and, as Crew tells us, there were in this experimental series 29 rats which went through the ordeal without making a single error, whereas in the control series there were only 10 such rats; and, as the table shows, each of the last 8 generations ranges down to zero errors. It would be of interest to know the distribution of the 29 errorless rats and of others of small numbers of errors throughout the experimental series. But the excess of these errorless rats in the experimental group over those of the control group is large enough to require explanation and to suggest a positive Lamarckian effect in spite of the defects of the procedure.

REFERENCES

(1) McDOUGALL, W. (1927). "An experiment for the testing of the hypothesis of Lamarck." This *Journal*, XVII, 267.
(2) —— (1930). "Second report on a Lamarckian experiment." This *Journal*, XX, 201.
(3) RHINE, J. B. and McDOUGALL, W. (1933). "Third report on a Lamarckian experiment." This *Journal*, XXIV, 213.
(4) CREW, F. A. E. (1936). "A repetition of McDougall's Lamarckian experiment." *J. Genet.* XXXIII, 61.
(5) McCULLOCH, T. L. and PRATT, J. G. (1934). "A study of the pre-solution period in weight discrimination by white rats." *J. Comp. Psychol.* XVIII.

[1] Here I may suggest again that just as there is in the individual in course of training a critical point, a degree of cumulative effect which must be exceeded if the effect is to be outwardly manifested in behaviour, so also it may be with the stock. In this connexion compare the evidence cited from Jennings in Part III of this article on the time factor in Lamarckian Transmission in unicellular organisms.

(*Manuscript received* 20 *July* 1937)

CRITICAL NOTICE

By M. D. VERNON

The Backward Child. By Cyril Burt. London; University of London Press, Ltd. 1937. Pp. xx+694. 20*s.* net.

In the whole of England and Wales there are, Prof. Burt calculates, nearly half a million children between the ages of 7 and 14 who, though not mentally deficient, are 'backward', or educationally retarded, to such an extent that their presence in the ordinary schools constitutes a real difficulty. They are, in fact, in need of special educational provision. In the county of London alone there are at least 40,000. For the sake of classification, Prof. Burt defines as 'backward' those whose educational attainments are between 70 and 85 % of the normal; who, without being mentally deficient, are yet unable in the middle of their school career to do the work of the class next below that which is normal for their age. Thus they must either work with children much younger than themselves, or fall hopelessly behind those of their own age. In either case they profit but little by their education; they are a nuisance to their teachers; and it is from their ranks that "the bulk of our criminals, paupers and ne'er-do-wells are eventually drawn". Clearly they constitute a social and educational problem of the first magnitude.

Prof. Burt, as Psychologist to the Education Department of the L.C.C., was brought face to face with this problem; and this book describes in great detail, and with the utmost lucidity and interest, his attempts, first, to understand it, and secondly, to recommend measures for dealing with it. Three steps were necessary: first, to find the causes of backwardness; second, to determine the extent to which each cause is remediable; third, to suggest the most hopeful line of treatment. Since it is chiefly to the practical solution of the problem that Prof. Burt has directed his attention, the book is intended primarily for the guidance of teachers, school medical officers, etc., and only secondarily for the treatment of theoretical psychological questions. But to the discussion of the latter is devoted a wealth of footnotes which, interesting in themselves, rather break the flow of the main argument. The psychologist is tempted to ask that Prof. Burt may some day give a whole volume to the ex-

haustive treatment of these questions, some of which are suggested below.

As regards the diagnosis of the causes of backwardness: Prof. Burt's method, which is similar to the one he employed in the investigations described in *The Young Delinquent*, was the analysis of the main and contributory factors producing backwardness in a sample of 400 children from London schools; and the incidence of these factors was checked against their incidence among a control group of 400 normal children of the same ages and types of environment. (A further analysis of the same kind was also carried out in Birmingham.) As might be expected, the causation was found to be multiple. In over half the cases there appeared to be inborn (sometimes hereditary) and irremediable 'dullness'—subnormality of general intelligence not amounting to mental deficiency. But frequently the effects of this natural dullness were aggravated by poor physique and health (especially by the after-effects of malnutrition and by chronic catarrh), by temperamental defects, and by poor surroundings and home background. Where no innate general intellectual defect was apparent, backwardness appeared to be due to specific defects of perception, memory or reasoning; or to instability of temperament—sometimes with real emotional maladjustment; or, very frequently, to specific physical defects of sight, hearing or speech, or left-handedness. Such troubles need not of themselves produce backwardness, as was shown by the frequency of their incidence in the control group. But they were more numerous in the experimental group, so that it appeared that backwardness might be produced by a kind of mutual reinforcement of unfavourable circumstances. But these difficulties, unlike inborn intellectual defect, were very frequently amenable to treatment; their worst effects, at least, were remediable.

The main characteristic of such treatment is its adaptation to individual needs. The child requires, in addition to actual remedial treatment for specific defect, a general modification of the normal educational system. He should be taught in a small class, if possible of backward children only, by specially qualified teachers, working with concrete objects and familiar situations rather than with abstract reasoning and generalized knowledge. Above all, the child must be dealt with as a particular individual. Prof. Burt stresses throughout that he should be regarded, not as a case of a particular type of defect, but as a whole individual, to whose backwardness many different factors have contributed. Even such relatively specific causes as speech defects and left-handedness (which Prof. Burt describes in great detail) are shown to be

far more complex in nature and variable in origin than is commonly supposed. Thus each child suffering from such a defect may present a different aetiology and require a modified form of treatment. But Prof. Burt's extremely analytical treatment is perhaps rather inclined to defeat his own object—namely, the re-synthesis of contributory causes into a unified clinical picture. We miss the illustrative case histories which were so admirable a feature of *The Young Delinquent*—here, alas! omitted, since the book is already very long.

But in addition one is sometimes inclined to feel that he has unduly multiplied the number of possible causes of backwardness—general and specific cognitive factors, emotional factors, moral factors—as a result of his adherence to McDougall's instinct theory and Spearman's factor theory. Is it not possible that many of these are but the apparent symptoms set up by one underlying cause—a fundamental lack of drive or motivation—and that many cases of backwardness are best regarded in this light? Thus, inability to remember, to form habits, to attend, to concentrate, are each treated under a separate heading. But the author himself gives an indication that lack of interest is the predisposing factor in these inabilities, though he nowhere defines or analyses interest. In the last chapter is given a brief discussion of methods of arousing interest and of supplying motivation. But these appear to touch only the fringe of a problem which seems quite fundamental to the treatment of the backward child. He may be already limited by poor intellectual and physical powers. But the essential difficulty lies in encouraging him to use in school such powers as he possesses. Why *should* such a child learn what is taught him in school? A more alert and intelligent child is eager to find out and to learn; not so the dull, the sick, or the child preoccupied with emotional difficulties. In the upper social classes, illiteracy and complete lack of culture are regarded as a social stigma. In the lower social classes (and over 20 per cent. of the school children in the worst London slums are backward as against 1 per cent. or less in the best districts), there is very little social incentive to acquire even the rudiments of knowledge, though the prestige and enthusiasm of the teacher may do much. Is it not then essential to indicate as a major cause of backwardness the lack of any ascertainable motive for producing learning in school?

All this must be taken, not as minimizing the value of Prof. Burt's treatment of backwardness, but as indicating the existence of a fundamental educational and social problem—the unsuitability of so much educational attainment to a large section of the population, as demonstrated in the existence of the 'backward child'. But, we may hope that

now he has so ably explored the overt causes of backwardness, he may some day turn his attention to the problems of interest and motivation, and their relation to innate impulsive tendencies; and devote to their study the method of combined observation and analysis which he has already employed to such good effect. Might not even the apparently irremediable limitation of general intelligence be found to be affected by such factors?

CRITICAL NOTICE

By E. FARMER

Accidents and their Prevention. By H. M. Vernon. Cambridge University Press. 1937. Pp. ix+335. 15*s.* net.

Recherches Expérimentales sur les Causes Psychologiques des Accidents du Travail. Par J. M. Lahy et S. Korngold. Paris. Conservatoire National des Arts et Métiers. 1937. Pp. 73. 20 fr.

These two books each contribute something to the serious study of accidents; the one by collecting relevant information, the other presenting an original piece of research work. Dr Vernon's book is probably the best textbook yet published on accidents. He presents clearly the results of much experimental work and also statistical data bearing on the study of certain aspects of accident causation. He deals with the frequency and severity rates of accidents under different environmental conditions and offers many valuable suggestions on accident prevention. He has collected a great deal of useful statistical information which he uses as the basis of his suggestions about accident prevention. In this way he keeps closely in contact with reality, so that his suggestions are not mere opinions but conclusions which follow from carefully considered data. The book as a whole is an excellent example of how statistical information can be used constructively as a guide to both theory and action and not remain a mere chronicle of figures.

The conclusions he draws are for the most part warranted by the available data, but doubts may be entertained whether some of these data really prove the conclusions he reaches. For instance, while he admits that most of the data showing that the consumption of alcohol increases accident incidence refer to excessive drinking, yet there is a general inference in the book that even a moderate consumption of alcohol increases accident liability. This may be so, but the data—both direct and indirect—at present available are not such as to warrant a definite conclusion. Dr Vernon gives an account of the work that has been done on accident proneness and shows that it is possible by means of psychological tests partially to detect beforehand those who subsequently have an unduly high accident rate. He thinks, however, that the progress made in this direction "has not been at all striking", and that

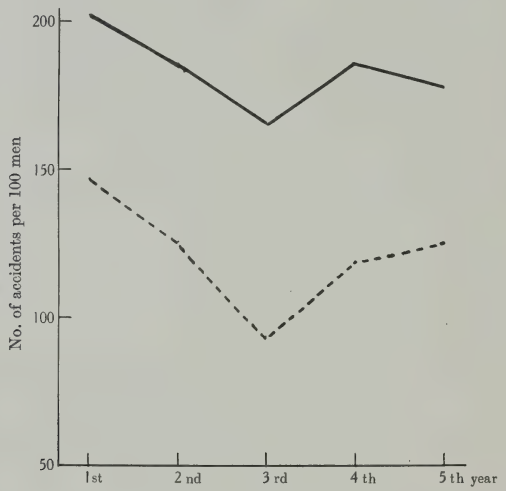
for the present it will be necessary to use "the rough and ready method of trial and error" for eliminating from a dangerous occupation those specially prone to accidents. The results obtained by those who have worked on accident proneness are not 'striking', if by this is meant that they have reached finality, but even in their present stage of development psychological tests are more prognostic of future accident incidence than the method of trial and error advocated by Dr Vernon.

The trial and error method, by which is meant removing from a department those with a high accident rate, depends on the stability of accident susceptibility between one period and another. Selecting workers with a low degree of accident proneness for a dangerous occupation depends on the predictive reliability of the tests used. Chambers and I (*Industrial Health Research Board*, Report No. 68. London: H.M. Stationery Office, 1933) have shown that both methods reduce the future accident rate, but that the latter does so more than the former, and that both combined reduce it still more. As far as I know we are the only investigators who have directly compared the relative value of the two methods, so that the available data are contrary to Dr Vernon's conclusions. We have data derived from a study of motor accidents supporting our original conclusion, but these are not yet published. It is unfortunate therefore that on this rather important point Dr Vernon's conclusions should be at variance with the data.

M. Lahy and Mme Korngold present in a clear form an original research on accident proneness among railway employees. Their general conclusions are that only those tests which impose a task on the subject, the timing of which is not under his control, have diagnostic value for accident proneness. In simple sensori-motor tests in which the subject can determine his own time of response, the performance of the accident and non-accident groups of subjects was equal.

Their findings are in accordance with those Chambers and I have already published. Certain of our tests that had proved prognostic of accident proneness have been used by M. Lahy and Mme Korngold and have yielded similar results to ours. The consistency between these two independent investigations adds to their value by reducing the likelihood of the results being obtained by chance. The authors draw the conclusion that those who are unable to deal with the accident situation are not necessarily those with slow reactions but those who are unable to deal with a situation imposed from without and which must be responded to in an arbitrarily determined time. Failure to respond to such a situation may be due to emotional factors affecting speed of co-ordinated reaction

at a complex level. At present only complex sensori-motor tests have consistently distinguished between those specially prone to accidents and others. It does not, however, follow that the real differentiation is at the sensori-motor level. Emotional maladjustment may be the cause of failure to respond to a complex sensori-motor situation, although the failure itself is only measured at the sensori-motor level. Sensori-motor tests may measure an effect common to their performance and to the accident situation, but the cause of this effect may be affective in nature.



Accident rate of a group of motor drivers over a five-year period.
Comparison of the rates of those who did well on the tests with those who did badly.

The fact that simple sensori-motor tests do not distinguish between those specially prone to accidents and others may be because the real differentiation is not sensori-motor but affective. This emotional reaction, in normal people, is only called forth in a complex situation involving a highly controlled response.

In referring to accident proneness the authors misunderstand the meaning Chambers and I attached to it when we first defined it. They think that since such factors as experience and age affect accident rate, the doctrine that proneness is the dominant factor in accident incidence

must to some extent be modified. It appears to us to be a more useful concept to regard accident proneness as an individual quality of relative permanence and to use accident liability as the sum total of all the factors determining accident causation. Liability, which can certainly be modified, includes proneness in its connotation, but the lowering of liability may not involve any change in proneness.

An example from some unpublished work of ours may make this clear. It will be seen in the figure (p. 352) that accident rate is lowered by experience both for those who are specially prone and for those who have a low degree of proneness, as measured by our tests, but the lines of improvement remain parallel. This means that the liability of both groups has been lowered by experience but that the liability of the specially accident prone, even after five years' experience, is higher than that of the others, on account of their greater proneness. This distinction between proneness and liability makes discussion easier and avoids certain invalid criticisms that have sometimes been made against the concept of accident proneness. Because an individual is accident prone it does not necessarily follow that he will have an undue number of accidents. If he is specially cautious or specially trained he may reduce his liability, which is the really important thing, although he will never bring it to the level of those who have a low degree of proneness and are equally cautious and well trained.

M. Lahy and Mme Korngold have used a large number of tests with 500 subjects and their results have been carefully analysed and clearly presented. The care with which their work has been done makes it an important original contribution to the study of accident proneness.

PUBLICATIONS RECENTLY RECEIVED

In the Realm of Mind. By C. S. Myers. Cambridge University Press. 1937. Pp. 251. 7*s.* 6*d.* net.

In this book Dr Myers has collected together nine lectures which he has given during the last few years to various audiences. All of them have been modified and put into literary form. At first sight it looks as if they must appeal to a variety of interests, and as if only the professional psychologist, if even he, will be interested enough to read the whole volume through from beginning to end. Two of the topics, on The Choice of a Career, and on Accidents, have direct industrial significance; one, on Musical Appreciation, is more a matter of pure psychology on the aesthetic side; two, on Social Psychology and Internationalism, are addressed to the student of current social and political life; two, on the bearing of psychology upon other sciences and on Medical Education, are particularly designed to interest scientists who are not themselves working in psychology; and two, on Mind and Body and on the Nature of Mind, have large philosophical implications. Yet it is difficult to imagine that anybody, beginning to read anywhere in this book, will fail to study all of the rest. There is a consistent biological approach all through, and in spite of the diversity of topics the essays display a unity of outlook and of exposition which must leave every reader the richer in knowledge and in appreciation of the contributions which psychology must be allowed to make towards human well-being. There is no sensationalism in the book. There is no indication that anybody with a slight smattering of psychological training can solve easily the abundant problems of civilization, but no topic is discussed without the most scrupulous regard for the facts that are most relevant to the issues involved. These essays could have been written by nobody without a profound and very extensive knowledge of modern psychology. They are written with the most remarkable lucidity and with no small amount of charm. They are unusual in that they can be read with profit by the psychologist who has years of experience behind him and also by the student whose years of experience are still to come.

General and Social Psychology. A Text Book for Students of the Social Sciences. By Robert H. Thouless. London: University Tutorial Press. 1937. Pp. xii+522. Price 8*s.* 6*d.* net.

This book is a revised and extended edition of Dr Thouless's well-known *Social Psychology*, which was first published in 1925. The new edition contains much extra material (it is nearly 150 pages longer than the earlier edition), and most of the original has been radically revised, so that the result, as the author says, is virtually a new book.

It is an extremely comprehensive textbook, covering most of the branches of psychological enquiry. The main stress is, of course, on problems of motivation, but the discussion of cognitive processes is adequate for the book's purpose. Freud, Jung, and Adler receive some attention, though there is no attempt at a detailed study of abnormal psychology. There are interesting sections on the psychology of morality, economic value, aesthetics, and of scientific and religious development. A new chapter on the use of statistical methods (a valuable introduction to the subject) precedes a much expanded treatment of mental testing. There is a new and stimulating discussion of the interaction of social groups and of the psychology of social classes. The book

is written in such a way that the student who does not wish to cover the whole field may omit some parts and yet find the rest intelligible.

Many of the problems with which Dr Thouless deals are, of necessity, problems on which there is not very much controlled evidence; but in discussing all of them Dr Thouless maintains an empirical and balanced outlook, and an unambiguous terminology, which should be of the greatest value in helping the student to think clearly.

The book can be recommended, not only as an excellent textbook for students of the social sciences, but also as a textbook which should be read carefully by students of general psychology.

Psychology and the Social Order. By J. F. Brown. New York and London: McGraw-Hill Publishing Co., Ltd. 1937. Pp. xiv+529. 21*s*.

This book differs from most books on social psychology in two important respects. First, it adopts as its theoretical foundation the topological psychology of Lewin. Secondly, the range of its problems owes much to the influence of Marx and his followers.

The effect of Marx on Dr Brown's conception of social psychology seems to be almost entirely beneficial. The fact that the book discusses such problems as the conflicts between social classes and the nature of a revolutionary situation brings it in contact with real problems of modern life which traditional social psychological textbooks have too often ignored. It is less easy to feel satisfied with the effect of the influence of Lewin. There are interesting topological diagrams to show social field structure in various situations such as depressions and revolutionary situations, but it is difficult to be sure that these diagrams really accomplish anything that could not equally well have been done without them. Topological methods seem also to be treated much more uncritically than they are by Lewin himself; what is to Lewin a useful tool seems to be treated by Brown as a system of dogma. When Dr Brown contrasts 'class theory' with 'field theory' he seems to fail to realize that the limitations as an intellectual tool of the method of making sharply divided classifications are assumed as a commonplace in all modern thinking at any rate in the biological sciences, and are not a matter peculiar to any one school. Also he seems to forget that the method of classification remains basic in all science, however much we may find it necessary to transcend its limitations. His sharp separation of class theory and field theory occasionally lands the author in absurdity. On page 40, for example, we learn that class theory analysis allows dichotomies whereas that of field theory allows no dichotomies. From this it follows that the dichotomy of 'class theory' and 'field theory' results from a lapse of Dr Brown into the ways of thinking characteristic of class theory. R.H.T.

Gesetze des Sehens. Von Wolfgang Metzger. Frankfurt a.M.: Kramer and Co. 1936. Pp. 172, 194 illustrations. R.M. 4.50.

It is remarkable that the very entertaining facts of visual perception investigated in psychological laboratories during the last fifteen or twenty years have not hitherto been made the subject of a popular book. It seems that, if such a book were copiously provided with illustrations so that its readers might make many of the basic observations and experiments on themselves, it could not fail to rouse wide interest. Prof. Metzger's book admirably fulfils this condition. There are pictures illustrating animal camouflage and most of the important Gestalt experiments. The one fact that is likely to limit its popular appeal in this country is that it is in German. It is to be hoped that this limitation will before long be removed by translation. Even now it is to be confidently recommended to any student of psychology who would like to have, without much effort, a preliminary acquaintance with Gestalt work on visual perception before he starts the more arduous task of reading Koffka's *Principles of Gestalt Psychology*.

Twins: A Study of Heredity and Environment. By H. H. Newman, F. N. Freeman and K. S. Holzinger. Chicago: University of Chicago Press. 1937. Pp. xvi+369. 18*s.* net.

The first part of the book contains a comparison between fifty pairs of identical and fifty pairs of like-sexed fraternal twins, each pair having been reared together. Ten criteria were used in diagnosing identical twins. All were subjected to a number of physical measurements and mental tests. In most of the traits measured the correlation coefficients were higher for identicals than for fraternals, a result that confirms previous studies of the same subject. In the second part of the book material is collected concerning nineteen pairs of identical twins reared apart for periods ranging from 11 to 53 years. Seventeen of the nineteen pairs were separated when under 2 years of age. The distance of separation, both geographical and social, varied considerably with the different pairs. It is this part of the work which makes the study unique and unlikely to be repeated for many years.

When the amount of difference between the separated twins was compared with the estimated amount of difference in their environments it was found that the highest correlation was between schooling and educational achievement (+0·91), a large but otherwise not very surprising result. Other correlations were between schooling and intelligence (+0·46 to +0·79), between social environment and intelligence (+0·32 to +0·53) and between physical environment and weight (+0·60). Seven of the nineteen pairs differed by ten points or more in Binet I.Q.: one pair showed a difference of nearly 4 years in mental age, two of 3 years, three of 2 years, and four of between 1 and 2 years. A comparison between the average differences of the identical twins reared together with that of the identical twins reared apart showed that the differences were significantly greater in the case of the identicals reared apart in weight, intelligence and school achievement. But in height, head measures and scores on the Woodworth-Mathews neurotic inventory the differences were not significantly greater. Thus some characteristics are more susceptible to environmental influence than others. Comparing the separated identical twins with the fraternal twins it was found that in some of the physical characteristics, notably height and head measurements, separated identicals were more alike than fraternals, and approximately as much alike as unseparated identicals. In weight, intelligence and educational achievement, however, separated identicals were on the whole as different as fraternals. Comparing the two sets of differences (between identicals reared together with identicals reared apart on the one hand, and fraternals reared together with identicals reared apart on the other) striking differences were found. The authors were led to the conclusion that "if the environment differs greatly as compared with heredity, the share of environment in determining traits which are susceptible to environmental influence is large. If, on the other hand, there is a large genetic difference and small environmental difference, the share of heredity is relatively large. This is what makes the solution of the question as to the relative share of the two sets of factors indeterminate." At the same time some of the differences in environment between the separated identicals were sufficient to produce differences in weight, ability and behaviour large enough to overshadow the genetic differences which occur between siblings. J.M.B.

Child at Play. By Marjorie Thorburn. London: George Allen and Unwin, Ltd. 1937. Pp. 187. 6*s.* net.

Mrs Thorburn kept a very detailed record of the play activities of her daughter Mary for the first four years of the child's life, and in this book a fairly full account, with some comments, is given. A general account of play development during the whole of the period observed is first presented, and this is followed by a much more detailed record of the period extending from 3 years 2 months to 3 years 5 months. Finally there are notes on spontaneous playthings, when and how often they were used, and

on the uses that may be made of records such as the one here presented towards a fuller understanding of the development of the mental life of the child. As a valuable source and reference book this volume must take a high place. The records were obviously most carefully kept and they are lucidly presented. No attempt is made to push any general conclusions beyond the limits to which they will validly reach; and if it stimulates students and parents to make many more observations of the same kind it will have performed a great service to child psychology.

Understanding our Children. By E. E. Read Mumford. London: Longmans, Green and Co. 1937. Pp. x+233. 3*s.* 6*d.* net.

This book is intended to replace Mrs Mumford's earlier volume called *The Dawn of Character in the Mind of the Child*, which, having had a very large circulation, is now out of print. It is an excellent book of the general kind, constantly illustrated by concrete discussion and uniformly sensible. Provided that the parent or teacher will recognize that his or her own child is likely to be different from every other, and will take the advice offered in the general sense in which it is meant, the book should do nothing but good. Most of the difficulties that confront the growing child are discussed with sympathy, and though often the specifically psychological statements made are over-dogmatic and emphatic the book as a whole must certainly help, as its author designs, towards a better understanding of child-life.

Modern Psychology and Education. By Mary Sturt and Ellen C. Oakden. Eighth Edition. London: Kegan Paul, Trench, Trubner and Co., Ltd. 1937. Pp. xi+307. 7*s.* 6*d* net.

This textbook, addressed to students in training colleges and in adult evening classes, has been extremely popular in this country. It differs from a good many textbooks of psychology in that it is full of original and lively illustration, mostly drawn from literary sources. It is very likely to stimulate the student to think about problems for himself and to find much material for their solution in his general reading and everyday experiences. Beyond a doubt this new and fully revised edition will take the book further upon its successful career.

Adolescent Psychology. By A. H. Arlitt. London: George Allen and Unwin, Ltd. 1937. Pp. 242. 6*s.* net.

This book was written for the purpose of giving sound psychological advice to those in charge of children. It certainly accomplishes its purpose. It is well balanced and free from fanaticism and should be of help to those who wish to understand the findings of modern psychology in relation to adolescent development.

The Children's Play Centre. By D. E. M. Gardner. (Contributions to Modern Education, edited by Susan Isaacs.) London: Methuen and Co. 1937. Pp. xv+143. 3*s.* 6*d.* net.

In a small cathedral town in England is a two-year Training College for teachers. The authorities wondered whether they could do something to bridge the gap that commonly exists between the study of psychological principles and the realization of practical problems. In 1931 they opened a Play Centre to which were invited, at that time, about twenty-four children from the poorer quarters of the town, of ages ranging from three to thirteen. This Play Centre was to be run mainly by the prospective teachers of the Training College.

In this book Miss Gardner gives a very excellent and comprehensive account of many of the things that happened. She describes how the available space was laid out and economically equipped, depicts typical evenings and types of play, tells of the rise and development of group activities among the children concerned, adds a few short studies of individual children, and shows in what ways the Centre helped both parents and the student teachers.

Miss Gardner is to be congratulated upon an extremely interesting book, and though it is perhaps intended mainly as a contribution towards the solution of certain problems of educational training, a great many of her descriptions of play activities and their organization have a psychological value of the most general character.

Psychology: the Changing Outlook. By Francis Aveling. (Changing World Library, No. 8.) London: Watts and Co. 1937. Pp. vi + 152. 2*s.* 6*d.* net.

The four main parts of this book are called Looking Backwards; Looking On; Some Applications of Psychology; and, Looking Forward. Looking Backwards gives a brief account of psychological origins and of the older intellectualist psychologies; Looking On describes in turn Noegenesis, Psychoanalysis, Gestalt and organismal psychology, Hormic psychology or purposivism, and Introspectionism and Behaviourism. In the summary an attempt is made to indicate the main lines of cleavage and agreement among current psychological theories. The applications dealt with are to education, to industry, to medicine and to criminology and law. As he looks forward Prof. Aveling is full of hope for the future, and is inclined, though with caution, to the view that a great many of the present difficulties may be solved by a more thorough and more experimental exploration of the phenomena of feeling and striving.

The book will serve its purpose most excellently. It can be put into the hands of any serious and reasonably good student with a certainty that it is likely to awaken his interest and to direct his attention to some real problems. It will indicate for him where he can follow up his studies most fruitfully. All the expositions, though they are notably brief, are both clear and adequate; and the selection of matters to be dealt with, which must have presented a most difficult problem, is exceedingly good. Already there were several extremely good books in the Changing World Library, and here is another one.

Leadership in a Free Society. By T. N. Whitehead. Oxford University Press. 1936. Pp. 259. Price 10*s.* 6*d.*

The author of this book is one of a group of workers at the Harvard Graduate School of Business Administration who are breaking new ground in Industrial Psychology by studying man's social reactions to the industrial environment. Dr Whitehead and his colleagues have already published the results of some of their investigations in social psychology in the industrial sphere. Their methods are definitely experimental and not confined to the generalizations that sometimes pass muster for social psychology. Some of these experiments are referred to in the book but in the main it deals with the general principles involved in the study of social reactions in industry.

The main theme is that modern industry by its rapid technological development has caused disintegration in the social life of those engaged in it. Social integration is possible for the industrial worker only within a system that permits of his voluntary co-operation towards an end desired not only by the group but by the individual. No specific solution is offered in this book but emphasis is laid on the necessity of understanding the social needs of the worker and of allowing for them in industry. If these social needs are disregarded by industry, then a cleavage arises within the worker's mind between his social organizations and his industrial organizations. The suggestion put forward by the author is that the latter should be an essential part of the former and not something opposed to it. He has in mind not the ordinary welfare schemes that are common to most large industrial concerns, but that industry should be organized in such a way as to fit naturally into a slowly changing social organization.

Psychology and the Promethean Will. By WILLIAM H. SHELDON. New York: Harper and Bros.; London: George Allen and Unwin, Ltd. 1936. Pp. x+265. 8*s.* 6*d.* net.

The author, a qualified medical man, is now professor of Psychology in Chicago Theological Seminary, so that it would seem that he is well qualified, on medical, psychological and religious grounds, for the study of the Promethean Will, i.e. of man's deepest wish to grow to full human stature, of his craving for knowledge of the proper orientation of life. This knowledge is not at present supplied by the doctor, or the psychologist, or the religious teacher, and Prof. Sheldon does not pretend to supply it either. But he has attempted a beginning, he has defined the target, and the outcome is a book which he might himself describe as "good fun". There is, as he says, "a religious sniffing in the air in many quarters" and many—psychologists, doctors, religious teachers and others—have, from their different individual points of view, outlined ways in which human development proceeds and stated the needs of our generation.

For the purpose of discussion the author takes particular 'panels' of the mind—the economic, political, sexual, religious and aesthetic panels—and discusses the conflicts in them with insight and understanding.

The book is addressed to no particular section of society and is written in language all can understand. To the 'waster' mind it will make no appeal; but the rest, whether or no they are aware of the Promethean urge, will find in its stimulating pages much food for thought. It is a book they will read more than once.

Dynamic Causes of Juvenile Crime. By NATHANIEL D. M. HIRSCH. Cambridge, Massachusetts: Sci-Art Publishers. 1937. Pp. 250.

In this book the author presents the results of some five years' clinical experience with delinquent children. He suggests that the causal factors in social structure are four—heredity, environment, accidental causation and genius—and it is on this basis that he considers the causes of juvenile delinquency, though he confines himself to the first three factors since he is not able to evaluate the effect of genius on juvenile delinquency. He shows that heredity, as he defines it, is the most potent factor in juvenile delinquency (60 %), that the environment contributes about 39 %, and that slightly over 1 % is due to accidental causation. The major causes are defective intelligence, instability, hyper-suggestibility, psycho-pathic personality, constitutional inferiority, ego-centricity, conflicts, endocrine disorders. Dr Hirsch ascribes greater importance to low intelligence than do many other investigators, but confirms the general high level of mechanical ability to be found among delinquents. He suggests that this indicates the urgent necessity for the provision of schools with an emphasis on manual and outdoor work. He also finds that there are no special difficulties to be faced by the only child, or by the oldest or youngest child in a family, as such. He also comes to the perhaps surprising conclusion that the effect of the Broken Home is negligible. The ill effects ascribed by others to the broken home are, he considers, due to the poor constitutional make-up of the parent or parents, which produces the poor milieu, which, in its turn, produces delinquency if, and probably only if, the child has a poor constitution similar to that of the parent. There is a chapter on enuretic delinquents, who suffer from an enuresis of unusual duration. Enuretic delinquents are, apparently, more incorrigible and prone to sex offences, but less inclined to larceny, than non-enuretic delinquents. The second half of the book is devoted to eight full individual case reports.

The work is an important contribution to the study of juvenile delinquency.

The Unknown Murderer. By THEODOR REIK. Translated by KATHARINE JONES. London: The Hogarth Press. 1936. Pp. 260. 12*s.* 6*d.* net.

In this volume Dr Reik discusses the reasons why murder exercises such a universal fascination. He brings a mass of evidence, ranging from a consideration of the methods used by primitive peoples in their attempts to find the murderer, and of the uses of oracles and ordeals in more advanced societies, to the most modern methods of detection, to bear on the problems involved, and shows the underlying connexions between them. From the same standpoint he considers the reasons why murderers (and other criminals) frequently leave clues behind them, often of so apparently senseless a nature that they lead to the criminal's detection. He also discusses the many psychological reasons why justice frequently miscarries.

Dr Reik does not claim to have settled these and other problems but his attack thereon, with its genetic and psycho-analytical approach, sets a new note in criminology which will be appreciated by those who are engaged in this field of research.

Inhibitions, Symptoms and Anxiety. By SIGMUND FREUD. Translated by ALIX STRACHEY. London: The Hogarth Press. 1936. Pp. 179. 6*s.* net.

In this work, which was published in Vienna under the title *Hemmung, Symptom und Angst* in 1926, Freud is principally concerned with the problems of anxiety and modifies his earlier views considerably. It is not that the cathectic energy of the repressed impulse is automatically turned into anxiety, or that symptoms are created to prevent anxiety from emerging. On the contrary anxiety is the danger-signal which (in a way Freud says he does not yet understand) causes, through the pleasure-unpleasure system, the inhibition of dangerous impulses; which symptoms are created in order to avoid danger situations whose approaches have been signalled by the emergence of anxiety. Anxiety first appears at birth. This sets the pattern for later anxieties, each stage of development of the ego having its own particular dangers, such as the danger of loss of the loved object in early childhood, the danger of castration in the phallic phase, etc.

The symptoms which appear in phobias, hysteria and obsessions are examined in the light of these findings.

A shorter account of the subject-matter of this book is to be found in Lecture XXXII of Freud's *New Introductory Lectures on Psycho-analysis* which was published in 1933, when an English translation also appeared.

Mind and Memory Training. By ERNEST E. WOOD. London: Pitman and Sons. 1937. Pp. viii+188. 5*s.* net.

The author of this book was the Principal of the Sind National College, Hyderabad. The volume is presented as the outcome of "over thirty years' experience and study of the memory systems of Europe and India, and it discloses hitherto unknown and interesting Eastern methods of mind training". It will be obvious from this that the treatise is definitely intended as a practical manual and one that will help the student to remember and to think. Prof. Wood claims to have discovered several "Roads of Thought" and these he advocates and illustrates with much ingenuity. Most of the advice turns upon the systematic utilization of particular forms of Association, and, as with all other mnemonic systems, it is maintained that that which is difficult to remember should be translated into terms of that which is, on the whole, and for most people, easier. Much also is written about concentration and various methods for improving it. Doubtless some students will find the practical devices advocated of some

value, though they will appear to others to burden recall with a great amount of indirect detail. To these the book must remain unconvincing, especially since the author's claims are not anywhere substantiated by rigorous experiments. The parts of the book which are said to deal with "unknown and Eastern methods" are very disappointing.

Love and Thought in Animals and Men. By SERGE VORONOFF. London: Methuen and Co. 1937. Pp. 138. 6*s.* net.

It is difficult to give a brief and adequate assessment of this book. Nearly two-thirds of it deals with mental processes assigned by the author to various animals. The method throughout is anecdotal, and the observational work involved does not appear to be either thorough or detailed. Dr Voronoff obviously has a great love for animals of all kinds and he writes of them with enthusiastic sympathy; but nothing very new and nothing that can be said to be properly demonstrated emerges. The second part, dealing with love and thought in men, also consists mainly of the description, very summary, of particular instances in which the love of man for woman is claimed to have exalted the man's capacity to think.

Successful Living. By BERAN WOLFE. London: Routledge and Co. 1937. Pp. 210. 7*s.* 6*d.* net.

This is another book of popular 'good advice'. It is prefaced by a diagram called The Basic Ego, which consists of a circle with a number of radiating 'spokes'. The four main spokes represent "the four cardinal points of social adjustment, occupational adjustment, sexual development, and the adjustment to the problem of leisure". The reader is invited to place himself somewhere or another on these radii, and then to consider his position in regard to numerous other radii all of which represent characteristics which have something to do with successful living. Most of the advice is gentle and pleasant; and it is apparently all regarded as so obvious as never to stand in need of proof.

BRITISH PSYCHOLOGICAL SOCIETY

REVENUE ACCOUNT AND BALANCE SHEET AS AT 30 SEPTEMBER 1937

Revenue Account for the year ended 30 September 1937

1936 £ s. d.	EXPENDITURE	£ s. d.	£ s. d.
170 0 0	*To* Rent, 55 Russell Square ...		170 0 0
252 0 0	Secretary's Salary		252 0 0
4 5 6	*Part-time assistance in Office*		
77 9 3	Printing and Duplicating ...		79 9 5
30 11 2	Hire of Halls		32 2 8
	Office and General Expenses:		
27 6 11	Caretaker, Lighting and Heating	30 5 8	
15 14 7	Stationery	10 10 11	
74 4 4	Postages and Telephone ...	70 0 9	
4 12 0	Fire etc. Insurance	4 12 0	
4 14 6	Subscriptions to Outside Associations	5 5 0	
15 15 0	Audit Fee 1936	10 10 0	
1 14 4	Repairs	1 18 1	
16 10 8	Teas	12 10 1	
4 17 5	Sundry Expenses	8 3 7	
165 9 9			153 16 1
	Journals Publication Expenses:		
239 3 6	General	166 9 3	
141 4 4	Medical	121 9 4	
98 0 0	Educational	89 5 10	
478 7 10			377 4 5
	Extended General Meeting 1937		22 3 8
9 1	Work for Committee on Human Mental Measurements ...		4 5 1
4 12 1	Midland Branch Expenses ...		17 5
	Northern Branch Expenses ...		2 12 1
10 13 5	Scottish Branch Expenses ...		9 12 5
22 3 9	Revision of Members' Register (completion)		18 14 3
1 10 0	Cheque Books and Bank Charges		1 0 2
50 0 0	*Publications Reserve*		
	Depreciation:		
	Library	64 15 8	
	Office Furniture and Fittings	19 15 0	
88 2 4			84 10 8
1361 14 2			1208 8 4
	Balance, being Surplus of Income over Expenditure		154 14 2
57 6 7			
£1419 0 9			£1363 2 6

1936 £ s. d.	INCOME	£ s. d.	£ s. d.	£ s. d.
	By Subscriptions 1937	1360 19 6		
	Subscriptions 1936	6 10 0		
		1367 9 6		
	Less Resignations and Removals ...	125 4 2		
1231 8 0			1242 5 4	
117 9 6	Interest on Investments		120 17 2	
70 3 3	*Recovery of Income Tax*			
£1419 0 9				£1363 2 6

Balance Sheet as at 30 September 1937

1936	LIABILITIES	£ s. d.	£ s. d.
£ s. d.			
2 7 0	Subscriptions paid in Advance		1 1 0
	Publications Reserve:		
122 0 0	Balance as per last Balance Sheet ...		122 0 0
	Revenue Account:		
	Balance as per last Balance Sheet ...	4610 19 11	
	Add Surplus of year	154 14 2	
4610 19 11			4765 14 1
£4735 6 11			£4888 15 1

1936	ASSETS	£ s. d.	£ s. d.
£ s. d.			
162 9 0	Cash at Barclays Bank Ltd.		238 8 6
3 5	Postage Stamps in hand		1 3 0
	Investments at Cost:		
	£300 5 % Conversion Loan 1944–64 ...	318 9 9	
	£3125. 11*s.* 4*d.* 3½ % Inscribed War Loan	3225 7 9	
3443 17 6			3543 17 6
	Note. Market Value of above Securities at 30 September 1937, £3487		
9 2	Sundry Debtors		1 9 3
126 10 6	Subscriptions in Arrear		111 6 0
	Office Furniture and Fittings:		
	Balance as per last Balance Sheet ...	197 10 0	
	Less Depreciation	19 15 0	
197 10 0			177 15 0
	Library:		
	Balance as per last Balance Sheet ...	596 0 0	
	Add Expenditure during year ...	51 15 8	
		647 15 8	
	Less Depreciation	64 15 8	
596 0 0			583 0 0
	Monograph Publications:		
	Balance as per last Balance Sheet ...	158 7 4	
	Add Grant made during year ...	50 0 0	
		208 7 4	
	Less Amounts repaid during year ...	26 11 6	
158 7 4			181 15 10
50 0 0	Shares in the British Journal of Educational Psychology Ltd. at Cost		50 0 0
£4735 6 11			£4888 15 1

I beg to report that I have examined the foregoing Accounts with the Books and vouchers of the British Psychological Society and that I have obtained all the information and explanations I have required. In my opinion the Balance Sheet is properly drawn up so as to exhibit a true and correct view of the state of the Society's affairs according to the best of my information and the explanations given to me and as shewn by the Books of the Society. The Securities appearing under the heading Investments have been verified by production of the Bankers' Certificate.

T. BARRETT
Incorporated Accountant
63 EASTCOTE LANE,
SOUTH HARROW
1 *November* 1937.

S. J. F. PHILPOTT
Hon. Treasurer.